2

REPORT RESUMES

ED 015 772 FS 000 222

TECHNIQUES FOR ASSESSING COGNITIVE AND SOCIAL ABILITIES OF CHILDREN AND PARENTS IN PROJECT HEAD START.

BY- HESS, ROBERT D. AND OTHERS
CHICAGO UNIV., ILL.

REPORT NUMBER OEO-519 PUB DATE JUL 66
EDRS PRICE MF-$0.75 HC-$5.64 139P.

DESCRIPTORS- *PREDICTIVE MEASUREMENT, *MEASUREMENT INSTRUMENTS, *EVALUATION, *PRESCHOOL CHILDREN, FAMILY ENVIRONMENT, CULTURALLY DISADVANTAGED, TEACHER CHARACTERISTICS, STUDENT ADJUSTMENT, ACHIEVEMENT TESTS, BEHAVIOR RATING SCALES, *PROGNOSTIC TESTS, ACADEMIC ABILITY, COGNITIVE ABILITY, HEADSTART, CHICAGO, STANFORD BINET, PSI, BI, DAP, READINESS CHKLST.,

IN ORDER TO DEVELOP AND EVALUATE MEASURING INSTRUMENTS FOR ASSESSING THE COGNITIVE CAPACITIES OF DISADVANTAGED CHILDREN, EXTENSIVE TESTING OF PRESCHOOL PUPILS AT 4 HEADSTART CENTERS IN CHICAGO WAS CONDUCTED. ACHIEVEMENT AND BEHAVIOR TESTS WERE ADMINISTERED DIRECTLY TO THE PUPILS. PUPILS' MOTHERS WERE INTERVIEWED AND TESTED TO OBTAIN INFORMATION ABOUT THE PUPILS' HOME ENVIRONMENT. OBSERVATION AND RATING EXERCISES BY TEACHERS AND OTHERS INDICATED SOME CHARACTERISTICS OF PUPIL ACHIEVEMENT AND BEHAVIOR. IT WAS HOPED THAT AS A RESULT OF THIS COMPREHENSIVE TESTING AND TEST-INSTRUMENT EVALUATION, A BATTERY OF EFFECTIVE INSTRUMENTS COULD BE IDENTIFIED THAT WOULD RELIABLY PREDICT PRESCHOOL CHILDREN'S SUBSEQUENT SCHOOL ACHIEVEMENT, EVALUATE THEIR SCHOOL READINESS, AND POINT OUT AREAS OF SPECIAL DISABILITY. IT WAS FOUND THAT INFORMATION ON THESE 3 AREAS COULD BE OBTAINED MOST RELIABLY BY MEASURING INTELLIGENCE AND ACHIEVEMENT BY (1) THE STANFORD-BINET, (2) THE PRESCHOOL INVENTORY, AND (3) THE DRAW-A-PERSON AND BY MEASURING BEHAVIOR AND ADJUSTMENT TO SCHOOL BY (1) CERTAIN ITEMS OF THE READINESS CHECKLIST, (2) CERTAIN ITEMS OF THE FACE SHEET OF THE STANFORD-BINET, AND (3) THE BEHAVIOR INVENTORY. IN A SUBSEQUENT STUDY COMPARING THE SCORES ON VARIOUS TESTS OF THESE HEADSTART CHILDREN WITH NON-HEADSTART CHILDREN WHEN IN KINDERGARTEN, NO DIFFERENCE WAS FOUND BETWEEN GROUPS IN ACADEMIC ACHIEVEMENT. (WD)

ED015772

THE UNIVERSITY OF CHICAGO
1892

PS 000222

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Techniques for Assessing Cognitive and Social Abilities of Children and Parents in Project Head Start

Report on Research Contract OEO-519
with the U. S. Office of Economic
Opportunity

Robert D. Hess - Principal Investigator

Richard Kramer
Diana Slaughter
Judith Torney
Carla Berry
Ethel Hull

University of Chicago

July, 1966

PS 000222

The research described in this progress report was conducted with the cooperation of community organizations and school district personnel in the Chicago area. We acknowledge our debt to these groups for their cooperation and assistance.

The research reported here represents the work of many people--testers, observers, data coders. The staff members who helped prepare the report were Mr. Richard Kramer, Miss Diana Slaughter, Dr. Judith Torney, Mrs. Ethel Hull and Dr. Carla Berry. Mr. Kramer and Miss Slaughter have also taken administrative responsibility for major aspects of the project since its beginning, and Dr. Torney has taken responsibility for major aspects of the analysis of the data. Miss Cynthia Schonberger has served as the project secretary during the majority of the project, and has typed this report.

We gratefully acknowledge the help and financial assistance given to this project by the Office of Economic Opportunity, Washington, D.C., under Contract Number OEO-519.

P.DH

INDEX OF CHAPTERS

	Page
I. Objectives and Design of Project	1
II. Assessment of the Cognitive Capacities of Head Start Children	6
III. Measurement of Social Behavior and Adjustment to the School	16
IV. Discussion of the Teacher as a Determinant in the Head Start Program	44
V. Naturalistic Observation as a Technique for Assessing the Behaviors of Head Start Children	50
VI. Assessment of the Cognitive and Socio-Emotional Home Environment of Lower-Class Head Start Children	69
VII. Head Start and Success in Kindergarten	89
VIII. Summary and Conclusions	106
Bibliography	
Appendix A - Instruments for Fall Retest	
Appendix B - Comparison of Performances on Three Administrations of the Pre-School Inventory By Item	
Appendix C - Behavior Form Observation Sheet	

INDEX OF TABLES

Page

I-1 Correlations of Cognitive Measures 13

III-1 Correlations of Readiness Checklist with other Behavioral Measures 21-22

III-2 Correlations of Readiness Checklist with Cognitive Measures 24

III-3 Items included in Summary Scores for Stanford-Binet Face Sheet Ratings Based on Both Teachers and Testers 26

III-4 Correlation of Binet Face Sheet Summary Scores with Behavior Inventory Summary Scores 28

III-5 Correlations of Binet Face Sheet Summary Scores with Cognitive Measures 29

III-6 Significant Correlations Between Selected Restrictive-Supportive Items [illegible] 31

III-7 Significant Correlations Between Selected Restrictive-Supportive Items and Behavior Inventory Summary Scores 33

III-8 Items Included in Summary Scores for the Behavior Inventory and their Factor Loadings 35-36

III-9 Behavior Inventory: Inter-Rater Reliability 38

III-10 Behavior Inventory Summary Scores: Distribution of Ratings by Teachers and Observers in Center B 39-40

III-11 Behavior Inventory: Changes in Summary Scores Over Time 42

III-12 Correlations Between Behavior Inventory Summary Scores and Cognitive Measures 43

IV-1 A Comparison of Product Moment Correlations of High and Low Teacher Groups with Predictive and Criteria Variables 49

V -1 Disagreements From Criterion Codes & Percent Agreement (for 7 coders coding 30 time samples) 59

Page

V-2 Distribution of Percentage Function Scores for Each Behavior Function 61

V-3 Number of Subjects Given each Percentage Level Score 63

V-4 Number of Children Given each Percentage Score for the 6 Types of Play 64

V-5 Speech 65

V-6 Number of Children Given each Mean Score For Each of the Four Observer Ratings 65

V-7 Significant (.02) Correlations of Naturalistic Observation Scores with I.Q., Achievement Test, Reading & Number Readiness and Report Card Scores 67

VI-1 Tabulation of Contacts with Parents 74

VI-2 Demographic Characteristics of the Two Head Start Centers 78

VI-3 Intercorrelation of Representitive Variables with Stanford-Binet Intelligence Scores of Working-Class Head Start Children 82-83

VI-4 Interrelation Between Preschool Achievement Test Gain Score (Pre and Post Head Start) & Selected Maternal Variables 85

VI-5 Intercorrelation of Mother-Teacher-Observer Behavior Inventory Factors 86

VII-1 Comparison During Kindergarten of Children with Head Start Experience and those without Head Start Experidnce - Age, School Performance and Cognitive Measures 91

VII-2 Rating of Behavior by Kindergarten Teachers-Readiness, Progress and General 93-94

VII-3 Rating of Behavior by Kindergarten Teachers-Speech, Participation, Achievement Motivation and Independence 94-96

VII-4 Summary Scores 98-99

VII-5 Correlation of Selected Variables from Summer Testing with Six Criteria of Performance in Kindergarten - Center A, Working Class Only 104-105

CHAPTER I

OBJECTIVES AND DESIGN OF PROJECT

A. Background and Statement of Objectives

This is the final report of a project designed to evaluate instruments which could be used to assess the needs and abilities of Head Start Children. The project had these objectives:

1. To evaluate the instruments distributed for research purposes by the Head Start Research office in Washington, D.C.
2. To develop or select additional instruments and techniques for assessing cognitive, social and emotional capabilities of preschool children.
3. To study the effects of input features of the Head Start program--the role and effectiveness of teachers, aides and volunteers in the project.
4. To study the impact of maternal behavior on social, emotional and cognitive performance of children in the Head Start program.
5. To contribute to the national pool of data on the characteristics and abilities of Head Start children.
6. To examine the usefulness of instruments for predicting school performance in preschool children.
7. To compare the performance of children in Head Start programs conducted under widely divergent philosophies of preschool education.

Underlying this project is the need for more adequate techniques for measuring behavior in preschool children in both privileged and underprivileged sectors of the society. The development of better data gathering apparatus has long-range implications for basic research and theory as well as for developing educational programs. Head Start offers the field of early childhood education a unique opportunity to study behavior of large numbers of children. The ultimate value of the data collected obviously will reflect the adequacy of the measuring devices.

Although this project utilizes populations from four separate Head Start Centers in the summer of 1965, its primary purpose was to evaluate instruments rather than programs. The choice of several locations with different curricula and philosophies of early education was designed as much to provide a wide variety of educational climates in which to work as to permit a comparison of results from the several Centers. In evaluating instruments assessing cognitive and emotional assets of preschool children, we are paying special attention to relationships between cognitive and social-emotional behavior. These two dimensions of behavior are not clearly differentiated in the young child, and both play important roles in school success.

We are placing considerable emphasis upon the development of data-gathering instruments for the use with parents of the children in the program. The interviews we have taken with Head Start mothers are thus aimed toward the refinement of interviewing assessment techniques and toward determining the relationship between maternal responses and the behavior of the child.

Although the goal of our project is to develop instruments for assessing and evaluating cognitive, emotional and environmental variables, our over-all follow-up design has allowed us to make some evaluative statements regarding the summer's Head Start programs.

B. Design of the Study

1. Description of Research Population

a) Centers

The Head Start centers in which the research was conducted varied in type of administration, teachers and student population. Our primary targets were Centers A and B, in which we duplicated a design which included naturalistic observations, lengthy interviews with parents and many testing and rating instruments In two other Centers, Centers C and D, we added to the sample requested by the national office on the Pre-school Inventory and obtained copies of all other instruments that were filled out in accordance with the design set up by the Research Director of Head Start in Washington. The Centers may be described as follow

1) Center A serviced a population of 126 Negro and 26 white children in a predominantly middle to upper-middle class suburb of Chicago. The building in which the program was held is an elementary school in the community. The teaching staff were all professional nursery school, kindergarten, or first grade teachers. By and large, teachers aides also had some professional experience. Volunteers were housewives from the community, (some with teaching experience) and high school students also from the community. This Center was chosen for its record of excellent cooperation in research endeavors and because it offered an example of a well run program sponsored by an outstanding school district. The instruments which had been tested in the summer were readministered to a sub-sample of our Head Start population in the fall as well as being given for the first time to a group of children enrolled in the same kindergarten classes who had not had experience in Head Start. This fall retest took place only in Center A where the concentration of post-Head Start children in three schools as well as the cooperatior of school officials made a follow-up study practical. We also gathered some information from the school's records - the scores on a nationally standardized test of reading and number readiness given in the spring as well as the child's grades from his report card at the end of the fall semester.

2) Center B serviced a population of 104 Negro children from a central city slum area in Chicago. The program was housed in a small four room "community house" adjacent to a church in the community. Unlike Center A, this program was run not by school district personnel but by a community action program. While all the

teachers had had previous teaching experience, only two of them had previously taught pre-school children. The aides were predominantly mothers of the pre-community without previous teaching experience. This Center was chosen for the contrasting nature of its student population, administrative and teaching personnel and orientation with respect to Center A.

3) Center C served a population of 72 Negro and Puerto Rican children in another central city location in Chicago. This program was housed in a parochial school and administrated through a Montessori teacher training Center. The teachers were all professional; the aides and volunteers were all residents of the community. This Center was chosen because of its contrasting student and administrative population and basic orientation.

4) Center D served a population of 60 Negro children in a central city slum area of Chicago. Like Center B, this program was administered through a community action program and run by members of the community who served in both teaching and administrative roles. This Center was chosen as a contrast to Center B in terms of administrative and teaching personnel.

The number of techniques administered in Centers C and D was limited. For that reason the majority of the data in this report comes from Centers A and B.

b) Social Class

Although Head Start is intended to be primarily for children from backgrounds of low social status, in each center there were a proportion of children who were from middle class, not working class backgrounds. For the majority of analysis in this report only the children from working class backgrounds were included This included children from homes where the head of the household was a laborer, domestic servant, skilled or semi-skilled manual worker or service worker. It also included those where the family receives public assistance.

2. Instruments Used in the Study

Instruments designed for this project, standardized instruments, and instruments designed under the auspices of O.E.O. which were used by this project to gain data on the children, their teachers, and their mothers were these:

a) The Children:

1) Naturalistic Observation. These observations were designed to obtain an account of the child's behavior in the school setting.

2) Summary Annecdotal Reports. A summary report on each child observed was written by both the project observer and the classroom teacher.

3) Checklist Evaluation for Kindergarten Readiness. This was a Questionnaire filled out by the teacher in conjunction with her summary report. This report included

ratings of the degree to which the observer and/or teacher liked or disliked the child, as well as the observer's and/or teacher's prediction of school adaptation and achievement in the early primary years.

4) Draw a Whole Person Test. An instrument used to assess the cognitive status of the child.

5) Techniques Designed to Measure Impulsivity. These were:

(a) Draw a Circle and Draw a Line Slowly . This is a task in which the child is asked to draw both the circle and the line as slowly as he can.

(b) Impulsivity Test. In this task the child was asked to sit still as long as he was able. The test is terminated at the end of 30 seconds.

(c) Delayed Reward. In this task the child was asked whether he would rather have a small piece of candy now or a large piece tomorrow.

6) Egocentrism Test. A Piaget type of task, designed to test the child's ability to adopt the perspective of another.

7) Length Conservation Test. A Piaget type of task, designed to test the child ability to conserve the concept of length.

8) The Stanford Binet Intelligence Scale

9) Templin-Darley Screening and Diagnostic Tests of Articulation. This instrument consists of 50 items which discriminate between good and poor articulation in pre-school and kindergarten children.

10) The Pre-School Inventory. This inventory was designed to find out whether the child has acquired certain skills that are ordinarily observable in children by the time that they are five years old. This instrument was designed by Dr. Bettye M. Caldwell for Operation Head Start.

11) The Operation Head Start Behavior Inventory. An instrument developed under the auspices of OEO consisting of 50 items which describe various types of behavior. The teacher and/or observer is asked to rate each child on a four point scale. In addition, for purposes of comparison, teachers and observers were asked to rate the children on these same items on a seven point scale. The rationale for this additional rating was that a four point scale might not allow sufficient range to discriminate adequately between children.

12) The Psychological Screening Procedure. A checklist of symptoms and descriptions of behavior problems of children. Teachers and observers were asked to check the symptoms and descriptions which applied to each child.

b) The Teachers:

1) Guide for Reporting the Teaching Situation. This instrument was used by

the observers at the end of the program as a technique for recording their observation on the methods of teaching used by the teachers in the center on which they observed.

2) Pre-School Teacher Questionnaire. An instrument designed by OEO to obtain the teacher's background and attitudes toward the disadvantaged child.

3) Interview for Teachers, Aides and Volunteers. An interview designed to obtain the teacher's evaluation of the Head Start Program and her attitudes towards the children in her class. The questionnaire was administered to the teachers at the beginning of the program; the interview was administered to the teachers at the end of the program.

4) Operation Head Start Workers Attitude Scale. An instrument designed to obtain the teacher's attitude toward the disadvantaged child.

c) The Mothers:

1) Parent Interview. A semi-structured interview constructed to gain extensive information about the mother and about the child's home environment.

CHAPTER II

ASSESSMENT OF THE COGNITIVE CAPACITIES OF HEADSTART CHILDREN*

Abstract: Several tests of intelligence and achievement, as well as tasks to measure impulsivity, egocentrism of thought, ability to conserve length, and speech articulation were included in the battery of tests administered to Head Start children. From an analysis of the relation of these cognitive measures to each other we conclude that if one wishes to measure the most important aspect of cognitive status in pre-school children, one should measure intelligence (or achievement/information) using either a Stanford Binet, the Preschool Inventory (PAT), or, if time and facilities are lacking, a Draw-a-Man IQ. In spite of some reservations about the use of tests standardized on middle class populations with Head Start groups, these tests (like the Stanford-Binet) have proved to be the most useful in assessing cognitive capacities. The use of measures of egocentrism, length conservation, and articulation, may be useful for the assessment of particular difficulties (as in speech), but are not generally useful. The tests of impulsivity need more research before a definitive statement can be made.

A. Purpose

One of the primary goals of this project was to develop and evaluate instruments for assessing the cognitive capacities of children enrolled in Head Start programs. The purpose of this evaluation was to recommend a set of instruments for use with working class children which could be used to predict their subsequent school achievement, to evaluate school readiness, and to assess areas of special disability.

B. Description of Instruments & Selection of Variables for Analysis

A variety of cognitive assessments were employed, including some standardized tests, some instruments pilot-tested by other investigators, and other tests developed especially for this project.

1. Tests of Intelligence

For this research tests of intelligence were particularly important measures of cognitive status. By comparing IQ scores with performance on other cognitive tasks it may be possible to limit the number of questions necessary for assessing those areas of cognitive functioning that are most important in evaluating school readiness

---

* This chapter was prepared by Judith Torney

a) The Stanford Binet

The intelligence of an elementary school child, particularly as measured by the Stanford Binet, has been the single most widely used assessment of intellectual ability and predictor of school achievement (Sundberg, 1960). However, there has been question raised both about the accuracy of predictions of elementary school success based upon IQ measured during the preschool period (Cronbach, 1960) and about the appropriateness of the Stanford Binet (and other intelligence tests standardized on white middle class populations) for measuring the academic potential of children from less privileged backgrounds, (Davis, 1948). The Binet is also a particularly lengthy procedure requiring highly trained testers. A test equally as good in predicting school success but shorter and easier to administer would be a great advantage in large scale assessments of Head Start children.

The Stanford Binet was administered once during the Head Start period. The IQ obtained from the administration of this instrument was used as a major variable in all analysis. The major conclusions concerning the usefulness of this variable are presented in the chapter concerning the prediction of school achievement, since it was exceedingly important to determine whether this test has a unique role which cannot be filled by any other test. IQ has also been used as a validating criteria for other cognitive and behavior measures used in this study. At the time of testing the testers also filled out ratings of Factors Affecting Test Performance. These are discussed extensively in the chapter on behavior ratings where they are called Face Sheet Ratings.

The mean IQ of the total group of working class Head Start children tested was 90.78, with a standard deviation of 14.51 (N = 187). The difference between the mean IQ in Location A and Location B was not significant.

b) The Goodenough Draw-A-Person.

This test (referred to hereafter as the Draw-A-Man Test) ranks second only to the Stanford Binet in the frequency of its use for intellectual evaluation of children in this country (Sundberg, 1960). Validity studies have reported correlation ranging from .65 to .74 between the Draw A Man test and the Stanford Binet (Goodenough, 1926; Harris, 1963). The ease of scoring and administration of this test, its adaptability to group testing, and its lack of reliance on verbal functioning made it an obvious choice for inclusion. Five Draw A Man Tests were administered in both group and individual test situations during the summer and one in the fall retest.

It was decided to use the Draw A Man collected in the classroom during the fourth week of the summer as the major piece of information from this test to be examined. This decision was made because the largest number of children has been present for this testing and because the second testing lacks some of the novelty of the first without having allowed too much practice in the task. The mean of Draw A Man IQ for working class children (N = 119) was 73.02 with a standard deviation of 13.70. The correlation of this score with the Draw A Man IQ obtained in the fall retesting was .575 (N = 88).

2. Tests of Information and Achievement

One test of specific information (comparable to elementary school tests of achievement) was used by this project in an effort to delineate those areas of information that are important to consider in evaluating a child's readiness for school.

a) The Preschool Inventory (Preschool Achievement Test or PAT).

This instrument was designed by Dr. B. Caldwell on the assumption that teachers have certain expectations of the child's level of information as he enters school. The instrument covers many areas. The child is asked his name, address, and the name of his classmates. His grasp of concepts of color, time, and ordination is tested. He is questioned on his knowledge of what mothers, soldiers, doctors, do. Further, his ability to follow instructions is examined. This project has compared the result of Caldwell's preliminary form of this Inventory with the standard intelligence measures as well as to examine its predictive validity as a test of school readiness.

The entire set of items was administered to the Head Start group in both Centers A and B during the third week of testing and again during the 7th week to a subgroup in each location.

As a result of complaints by Head Start teachers and observers that the Preschool Inventory was too bulky an instrument to be administered effectively or to sustain the child's attention during the necessarily lengthy administration, it was decided to shorten the instrument for the retest program planned for the fall. At this time, distributions in the form of percent of children passing each item were available for 2/3 of th e children in Center A for the initial summer administration. (See Appendix B)

It was decided to include in the retest administration items from all sections of the original instrument where the initial percentage of children passing would be low enough to allow for future change, as well as a number of high percentage-pass items so that the less achieving children would not be discouraged by a barrage of questions, none of which they could answer. Our aims, then, were threefold; to produce an instrument of reasonable length for children of this age, to produce an instrument with a high enough ceiling to allow future change, and to provide a group of questions, hopefully from each substantive area, which would range widely in difficulty. We also tried, as much as was possible, to exclude any items which depended primarily upon experience with the object as opposed to knowledge of the concept involved.

As an example of our procedure for item-selection, the original protocol included a ten-question subset of items involving the identification of bodily parts. For our revised instrument, we pulled from this subset four items (What's this? (a) Finger, (b) Eye, (c) Elbow, and (d) Heel) with percents passing of 83.8, 97.1, 57.7, and 26.7 respectively. As can be seen, a range of difficulty is represented.

Forty-nine items were eventually selected for inclusion in the revised instrument (see Appendix A and B). All substantive areas on the original instrument were not, however, represented. The group of twelve items on the original test involving instruction-following ("Close your eyes," "Raise your hand," etc.) were not included, as better than 85 percent of the children passed each item. None of the fourteen items involving identification of colors were included, as color-identification was involved in other items in the revised instrument (for example; "Put the red car on the black box").

The four items on the original instrument involving naming or indicating directions ("What way does a saw go? An elevator?" etc.) were not included, as it was felt that successful completion of these items depended too heavily on specific experiences which the child may or may not have had, rather than upon correct identification of direction.

The eight items which were originally designed to indicate perception of authority figures as restrictive or supportive were not included in the reduced instrument; at that time a suitable coding system had not been worked out and consequently no statistical information was available. These items are briefly presented in Chapter III.

A forty-nine item revised instrument was administered in the fall retesting. For the purpose of obtaining comparable scores for all three administrations the subset of items which had been administered on the retest were scored using the responses given at Summer 1 and Summer 2. These will be referred to as Partial Item Set-Summer 1, Partial Item Set-Summer 2, etc.

The correlation of the Partial Scores, scoring only 49 items with the Total Score scoring all 152 items for the first summer testing (N = 169) was .949. A part-whole correlation of this magnitude, even when the number of items used was approximately one third of the total, indicates that the entire test is highly consistent and that it can be cut considerably (thereby simplifying testing procedures). It also suggests that the results reported here with this set of items would probably be highly similar to those reported by other investigators who may choose to use the revision and selection of PAT items recently copyrighted by Caldwell & Soulé (1966).

The correlations between the three testings (Summer 1, Summer 2, Fall Retest) using the Partial Item Set of 49 items ranged from .800 to .885. Information and achievement at the preschool level are highly consistent even across a four month period. In the analysis of correlates of these variables, Summer 1 Testing (Partial Item Set) and Fall Retesting (Partial Item Set) were utilized more extensive than the Summer 2 testing where the number of children who took the test was drastically reduced.

The mean PAT scores for Centers A and B were not significantly different either at Summer 1 or Summer 2 testing.

3. Tests of Concept Formation

While intelligence tests have considerable face validity as measures of cognitive ability, their use for evaluation of school readiness poses some serious problems. The work of Piaget and the research it has fostered point to two important considerations in evaluating school readiness. First, it has been suggested that mental age alone is not necessarily the only indicator of level of cognitive development. Kohlberg (1961) has suggested the concept of developmental age (more closely associated with chronological age than with mental age), which may be useful in evaluating the child's readiness for certain types of learning situations. Second, Piaget's analysis of the pre-operational period of development (ages 2-6) suggests that children of preschool age are at that point when they are beginning to be able to make adaptive use of the intelligence they have. Prior to this, the child, whatever his mental age, is largely unable to be critical of his own thought processes It can be inferred that at least some children are unable to make efficient use of the learning experiences in school. As Flavell has said, "It is no accident that the lower age limit in most Piaget experiments is about four years." (Flavell, 1963, p. 162).

a) Egocentrism.

This Piaget task, developed by Dr. L. Kohlberg, examines the child's ability to adopt the perspective of another. Its possible usefulness as an indicator of school readiness is reflected in Flavell's description of the egocentric child. "(He) feels neither compunction to justify his reasoning to others nor to look for possible contradictions in his logic. And causally related to this, he finds it exceedingly difficult to treat his own thought processes as an object of thought." (Flavell, 1963, p. 156). In other words, the egocentric child would be at a disadvantage in performing in a school situation which asks him to be critical of his own thought. This test was administered in both locations during the summer and in location A Fall Retesting.

The test of egocentrism, dichotomously coded, showed a nonsignificant correlation of .171 (N = 87) between the Summer and the Fall Retest scores. It was included in the correlational analysis, which will be summarized in the following section.

b) Length Conservation.

In his book The Child's Conception of Geometry, Piaget (1960) has studied the child's growing ability to conserve and measure such geometric entities as length, area, and volume. His work indicates clear stages in the growth of these abilities. The length conservation task used by this project was developed by Dr. L. Kohlberg. The scoring of the child's responses to this task allow for the assessment of the child's developmental age which may prove a valid measure of school readiness and as a predictor of subsequent achievement.

4. Tests of Impulsivity

It has been argued that the ability to inhibit motor movement should be functional for problem solving. The work of Maccoby, Dowley, Hagen, & Degerman, (1965), showed positive correlations ($r = .44$, $p < .01$) between Stanford Binet scores and performance on inhibition tasks. It is clear also that the ability to inhibit movement is pragmatically important in the school setting.

a) Draw-A-Line and Draw-A-Circle Slowly Tests.

These two tasks ask the child to draw as slowly as he can, delaying the completion of the task indefinitely. Scores are length of line (diameter of circle) and number of seconds spent drawing. This task was administered twice during the summer and in the fall retesting (Location A). The scores used were the means over two trials for each task.

All of the measures derived from the Draw a Line and Draw a Circle Slowly tasks were correlated using the Center A population of middle and working class children to determine which measure (length or time) and which figure (line or circle) should be utilized in further analysis. Time spent drawing lines showed considerably higher correlations across time than any of the other measures. Time Lines Summer 1 was correlated .390 with Time Lines Summer 2 and .494 with Time Lines Fall Retest. Time Lines Summer 2 was correlated .312 with Time Lines Fall Retest. All these correlations were significant at the .02 level or better. The corresponding correlations for Length Lines were -.077, -.035, and .194; for Time Circles .145, .150, and .198; for Diameter Circles .274, .242, and .126. (Each of these measures was the mean length or time averaged across two lines or circles). Drawing lines is probably an easier and more familiar task for preschool children and one where impulsivity, not manual ability is the major determinant of the ability to inhibit movement. Time Lines Summer 2 and Time Lines Retest were used as the major variables for measuring impulsivity in the more extensive analysis (Fn: Time Lines Summer 2 was chosen because it has been administered in both Centers A and B)

b) Sit Still Test

This task asks the child to remain seated without moving for 30 seconds. Scores are Latency of First Movement and Number of Total Movements. This task was administered once during the summer and at the fall retesting (Location A). The Number of Movements in the Sit Still task and the Latency of First Movement were not significantly correlated over the summer to fall retest period. Number of Movements was not used in other analysis. Latency of First Movement was included in some correlational analysis which will be reported in the following section.

c) Delayed Gratification.

In this test the child is offered the option of having a small piece of candy now or a big one tomorrow. Mischel (1961) has suggested that the capacity to delay gratification for large reward (the opposite of impulsivity) is particularly important in socialization and the development of social responsibility. Children were given this choice once during the summer testing. The choice of Delayed vs. Immediate Reward, a dichotomous variable, was included in some correlational analysis There was no consistent clustering of this variable with the other impulsivity measures. The results suggested that the child who is capable of delaying is more likely to be high verbal, but level of correlation was not high. The dichotomous nature of the variable may account in part for its low correlations.

5. Test of Articulation (Templin Darley).

The Templin Darley Test of Speech Articulation was used as a screening device in Center A during the summer for placement in special speech classes and these scores were utilized by this investigation. It is apparent that speech ability is an important prerequisite of verbal functioning which is influential in determining scores on many verbal tests including the Stanford Binet.

C. Inter-Correlations of Cognitive Variables

1. Intelligence Tests

a) Stanford Binet

The Stanford Binet was significantly correlated ($p < .02$) with every other cognitive variable except the Egocentrism Retest. It was correlated most highly with the PAT administered in Summer 1, .733. Its other correlations with Behavioral ratings and with success in Kindergarten are discussed in appropriate chapters.

b) Draw-A-Man:

The Draw a Man tests were correlated with the Binet .488 (Summer) and .485 (Ret almost as highly as the summer-retest correlation of the D-a-M with itself. The Draw-a-Man IQ was also significantly correlated with the majority of other cognitive variables (with the exception of the two egocentrism tests and the latency of first movement in the Sit Still Task). However, in all cases these correlations were lower than the corresponding correlations with the Stanford Binet. Although the Draw-a-Man is a measure of IQ, it is not equivalent to the Binet and it has certain drawback its retest correlation is not as high as could be hoped, and the Draw-a-Man Retest IQ is significantly correlated with chronologic al age, which should not be true of a score which has been normed for age as the IQ is. This test can be used as a measure of IQ when limitations of administration and time make it more efficient. Its predicting ability for school achievement will be discussed in Chapter VII.

2. Test of Achievement and Information:
Preschool Achievement Test (Preschool Inventory)

The PAT scores, both summer and retest, are significantly correlated with chronological age as would be expected of tests which are not normed by dividing by age as IQ scores are. This is important however in interpreting any changes which occur over time in PAT scores. If the PAT score shows a gain over the period from Head Start to fall retest, and children have higher scores in the fall this corresponds to an increase in the mental age component of the IQ. It would be expected that children would acquire new information in this period. Therefore an increase in PAT scores would not necessarily have the same meaning as an increase in IQ. An increase in IQ means that the child's mental age has increased more rapidly than his chronological age; an increase in PAT score means simply that the child's achievement has increased. There is no way to tell how great this increase is relative to his chronological age.

TABLE I-1

Correlations of Cognitive Measures[a]

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12
(1) Chron. Age												
(2) Stanf-Binet	-.102 (187)											
(3) Draw a Man Week 4	.165 (116)	.488* (116)										
(4) Draw a Man Retest	.286* (86)	.485* (86)	.575* (88)									
(5) PAT Partial Summer 1	.308* (160)	.733* (160)	.474* (105)	.568* (81)								
(6) PAT Partial Retest	.316* (87)	.684* (87)	.433* (89)	.520* (89)	.800* (82)							
(7) Egoc. Summer	-.044 (185)	-.268* (185)	-.179 (116)	.054 (86)	-.317* (158)	-.190 (87)						
(8) Egoc. Retest	-.002 (87)	-.209 (87)	-.111 (89)	-.187 (89)	-.198 (82)	-.275* (90)	.172 (87)					
(9) Timeline Summer	.120 (164)	.309* (164)	.233* (114)	.222 (85)	.296* (142)	.335* (86)	-.041 (163)	-.298* (86)				
(10) Timeline Retest	.240 (87)	.515* (87)	.359* (89)	.463* (89)	.544* (82)	.461* (90)	-.067 (87)	-.152 (90)	.250* (86)			
(11) Latency 1 Summer	.080 (166)	.308* (166)	.174 (99)	.010 (70)	.283* (139)	.086 (71)	-.276* (166)	.091 (71)	.194* (147)	.286* (71)		
(12) # Errors Articula.	-.174 (104)	-.273* (104)	-.265* (107)	-.395* (80)	-.320* (96)	-.431* (81)	.044 (104)	.058 (81)	-.142 (102)	-.390* (81)	-.270* (89)	

* Significant at the .02 level of probability

[a] These correlations were based on working class children only.

The test-retest correlation of the Partial Item Set from the PAT was .800, considerably better than the Draw-a-Man Test. From correlation of both administration of the PAT with the Binet (.733 and .684) it appears that the distinction in test content between achievement and intelligence is not clear. The Binet in fact uses a large number of information questions in assessing intelligence. The PAT is also highly correlated with the Draw-a-Man tests (.474, .433, .468, and .520). When testing time and personnel are limited, either the Draw a Man IQ or the PAT (either the Partial Item Set used by this project or the recent revision suggested by Caldwell) may be substituted for the Binet. This possibility will be discussed more fully in the chapter on prediction of school success.

3. Tests of Concept Formation

The correlations of chronological age with the two Piaget-type variables (Egocentrism Summer and Retest and Length Conservation) were not significant. This is particularly important because of the design of these variables as measures of development, not mental age. With the Egocentrism task, the lack of correlation may be in part the result of the dichotomous nature of the variable compounded by the relatively small range of chronological age of the group. The Length Conservation task score had a larger range, but this variable showed few meaningful correlations of any kind, indicating perhaps some problems with the scoring or administration.

The correlations of both Egocentrism Summer 1 and Conservation with Binet IQ was significant (though not high). The correlation of Egocentrism Summer 1 with Binet IQ was -.268, significant at the .02 level; the correlation of Egocentrism Retest with Binet IQ was -.209, not significant; the correlation of Length Conservation with Binet IQ was .229, significant at the .02 level. Rather than these tests measuring something about developmental age neglected by measures of mental age, it seems that these variables are weak measures of the same sort of ability that IQ and achievement tests measure. Findings in chapter VII indicate that these variables are of little importance in predicting success in school.

4. Tests of Impulsivity

The Mean Time of Lines is also highly correlated with all the intelligence and achievement measurements. In fact, a number of these correlations are higher than the Summer to Fall Retest correlations for Mean Time Lines (Fn: the discrepancy between this summer-retest correlation and that reported in the earlier section is due to the inclusion in the earlier correlation sample of both middle and working class children). The meaning and usefulness of this score are not wholly clear. The test-retest correlation is not as high as other intelligence measures, yet its correlation with Binet is very high. On the basis of this information it is not possible to recommend the use of this test as a substitute for the Binet. Nor are its correlations clear enough with other variables to suggest that this is a cognitive dimension of great importance independent of IQ. Further work needs to be done in this area.

Latency of First Movement in Sit Still task was significantly correlated with Binet IQ, PAT and Egocentrism but seems to contribute little to the total analysis.

5. Articulation

The number of errors on the articulation test was correlated with the measures of intelligence (Binet -.273, PAT Summer -.320). These correlations are of about the same magnitude as those reported between IQ and Egocentrism and Length Conservation and the Impulsivity Measures.

D. Summary

The variables may be grouped in the basis of this analysis into those which are clearcut tests of intelligence and achievement (Binet, PAT, Draw a Man IQ) and those which, though their correlations with these intelligence tests are significant are considerably lower and are not part of this cluster. (Egocentrism, Length Conservation, Articulation, Impulsivity). In order to determine whether this second grouping of tests would be useful for measuring aspects of cognitive performance other than intelligence factor, their correlations with selected other variables (Particularly Behavior Inventory Summary Scores) were examined. There was little evidence for the usefulness of these measures as unique predictors of other behavior in the schoolroom. (This evidence will be more clearly summarized in Chapter VII. In other words, as far as this analysis can determine, if one wishes to measure important aspects of cognitive status in pre-school children, one should first measure intelligence (or achievement/information) using either a Stanford Binet, Preschool Inventory, or (lacking time and facilities) a Draw a Man IQ. The use of the measures used here to assess egocentrism, length conservation, and articulation may be useful for the solution of particular problems or assessmen of particular difficulties (as in speech). The tests of impulsivity are in an inte mediate position and require more work. They are neither very good substitutes for tests of intelligence nor do they measure what appear to be crucial dimensions of cognitive functioning.

CHAPTER III

MEASUREMENT OF SOCIAL BEHAVIOR AND ADJUSTMENT TO THE SCHOOL

Abstract: Extensive analysis of four instruments (The Psychological Screening Procedure, the Readiness Checklist, the Face Sheet of the Stanford-Binet Intelligence Scale, and the Behavior Inventory) has been presented in this chapter. Partial analysis of the eight items on the original Preschool Inventory designed to measure perceptions of authority figures as restrictive or supportive has also been presented.

With heaviest emphasis on instrument assessment and reduction, we used Principal Component Factor Analysis to form clusters of homogeneous items on two instruments; the Behavior Inventory and the Stanford-Binet Face Sheet. We then summed ratings on each child for items within each item cluster to form Summary Scores, on which we based further analysis. Correlations between teacher and observer rating have been presented in support of inter-rater reliability on the Readiness Checklist and the Behavior Inventory. Tests for reliability could not be made on the Binet Face Sheet, as necessary data was not available. Correlational material has been presented in support both of relationships between behavioral instruments, and of relationships between behavioral and cognitive measures. Also presented are correlations between restrictive-supportive items on the Preschool Inventory and Summary Scores on the Behavior Inventory.

Recommendations: It is recommended that the Psych. Screening Procedure be excluded from future studies of this nature, as the items in this instrument are too extreme to apply to the majority of Head Start populations. If the aim were to find children who had any serious mental or physical disfunctioning it would be more efficient to ask the teacher whether any child appears to be seriously disturbed, since the majority of symptoms listed are of such an extreme nature that they would be quickly visible to an observer or teacher.

It is recommended that the twenty items included in the five Summary Scores (Agression, Verbal-Social Participation Timidity, Independence, and Achievement Motivation) for the Behavior Inventory be included in future studies, and that they be summed to form summary scores.

For the Face Sheet of the Stanford-Binet Intelligence Scale, Form L-M, it is recommended that the eleven items emerging in our factor analysis be used in future studies, and that they be summed to form the three summary scores, Achievement Motivation, Confidence in Ability, and Activity Level. These items can be used as rating scales for general, "long-term" behavior characteristics, as well as in test-specific situations.

It is recommended that seven of the Readiness Checklist items be included in future studies. These are: Kindergarten Readiness, Probable School Achievement, Probable School Adaptation, Appearance, Motor Coordination, Speech, and Engagement in Class. These items involve assessment of general school-oriented behaviors not included in

---

This chapter was prepared by Ethel Hull

instruments provided by the Office of Economic Opportunity.

No recommendations can at this point be made regarding the eight items from the Preschool Inventory involving perception of Authority figures as restrictive or supportive as exhaustive analysis has not yet been completed.

A. Introduction

Cooperativeness with other children, interest in talking and listening to others, ability to play without constant adult supervision, and energetic interest in new objects and experiences are among the social and emotional characteristics which foster adjustment and achievement in the early elementary school years. This study of Head Start attempted to assess these social and emotional characteristics by four types of rating instruments administered to testers, teachers and observers during the summer program, and to teachers during the fall retest program. The four instruments were the Psychological Screening Procedure (or Symptom Checklist), the Behavior Inventory, the Readiness Checklist, and the Face Sheet of the Stanford-Binet Intelligence Scale, Form L-M. The Screening Procedure and the Behavior Inventory were designed by the Office of Economic Opportunity to be used on a nation-wide basis; the Readiness Checklist was designed here at the Urban Child Center.

Analysis of the eight items on the original Preschool Inventory (PAT) (designed by Dr. Bettye C aldwell for the Office of Economic Opportunity) to measure perceptions of authority figures as restrictive or supportive has also been included in this section, as it was felt that these items not only apply more directly to the personality and behavioral areas than they do to the cognitive, but that they also involve direct responses by the child rather than ratings of the child by others.

B. Psychological Screening Procedure

The Psychological Screening Procedure or Symptom Checklist consisted of thirty-eight items concerning single behavioral, psychological or physical symptoms of maladjustment and nine items asking for mominations of children of a particular type (e.g. the withdrawn child). Each of these 47 items was to be answered Yes if a child had the symptom or No if he did not.

This instrument was considerably less successful than the Behavior Inventory (to be discussed subsequently) in providing useful information about the behavior, adjustment, and characteristics of the Head Start group studied. For thirty-nine of the forty-seven items fewer than 10% of the children were nominated (by either observers or teachers) as possessing the characteristic or symptom; for thirty-one of these items fewer than 5% were nominated. The items in this instrument list behaviors and symptoms which are very infrequent in the Head Start populations dealt with in this study (Fn: Approximately five percent of this group was refereed to a mental health or child guidance clinic and approximately seven percent were referred for medical attention.) The eight items where more than 10% of the group was nominated were the less extreme items in the instrument and are similar to the items included in the Behavior Rating. Furthermore, teachers and school administrators reacted negatively to the negative 'symptomatic' connotations of the items.

Because of the limited number of nominations on these characteristics, little analysis was done to determine the relationship of these items to other data. The dichotomous nature of the nominations also presented statistical problems. It might be that in other Head Start populations, perhaps in institutions or in more seriously deprived areas, the incidence of behavior problems would be great enough to warrant the use of this instrument to find children with a particular symptom. If the aim were to find all children who had any serious mental or physical disfunctioning it would probably be more efficient to ask the teacher whether any child appears to be seriously disturbed, since the majority of the symptoms listed are of such an extreme nature that they would be quickly visible to an observer or teacher. For most of the populations that will be assessed as part of the Head Start program, this instrument appears to be of limited usefulness.

3. Readiness Checklist

Designed at the Urban Child Center, the Readiness Checklist in its original form consisted of twelve items oriented toward readiness for and future progress in school. Children were rated by teachers, at the conclusion of the Head Start Program, on perceived Readiness for Kindergarten. This rating was made on a 5-point scale. Nine Ratings were then made on specific behavioral, emotional and experiential areas in which the child could be perceived as not being prepared for kindergarten. Two examples of these specific ratings are: "Not able to control his behavior," and "Not able to communicate his needs well enough to be easily understood by teacher." These ratings were made on a two-point yes-no scale indicating presence or absence of the problem in question.

Four additional ratings (here on a seven-point scale) were then made, by both teachers and observers, of their like and dislike of the child and for his probable adaptation and achievement during the early school years. Administration to both teachers and observers included children from Centers A and B.

All items from this instrument (with the exception of the like-dislike rating) were included in the fall retest sample; the dichotomous items were expanded to a three-point scale (very true, somewhat true, not true) to facilitate correlational analysis. Four additional items, originally a part of the naturalistic observations, were included in the retest. These were designed to measure general habits of Appearance, Motor Coordination, Speech, and Engagement in Class, and were rated on a seven-point scale.

1. Inter-rater reliability

Product-moment correlations, based only on working class children from Centers A and B, between teacher and observer ratings on Probable School Achievement and Adaptation were moderate though significant at better than the .02 level ($r = .416$, $N = 123$ for teacher vs. observer Achievement ratings; $r = .351$, $N = 115$ for teacher vs. observer Adaptation ratings).

It is believed that certain factors in one of the Head Start Centers operated to depress the magnitude of the correlations based upon the entire sample. This will be discussed in greater detail, with substantiating data, in Section F-3 of this Chapter.

2. Correlation of Readiness Checklist Items with Other Measures of Behavior

The summer administration to teachers of the Readiness Checklist correlated highly with Summary Scores of the Behavior Inventory, administered to teachers at the onset and at the conclusion of the Head Start program. The product-moment correlations between Readiness for Kindergarten and summary scores involving Aggression were high ($p = .02$ or better) and negative, with moderate negative correlations between Kindergarten Readiness and Timidity. Correlations between Readiness and behavioral scores for Verbal-Social Participation, Achievement Motivation and Independence were moderate to high and positive ($p = .05$ or better) (see Table III-1). An almost identical trend was found for correlations between the Readiness Checklist item for Achievement and summary scores for the Behavior Inventory. Correlations between Probable Kindergarten Adaptation followed a similar trend, with correlations of lower magnitude than for the variables cited above, especially in the case of the summary score involving aggressive behavior.

As might be expected, behavioral ratings and Readiness ratings by teachers showed generally higher correlations than did teacher-observer instrument comparisons, and vice-versa.

Behavior Inventory Summary Scores from the fall retest administration correlated highly ($p = .02$ or better) with the retest administration of the Readiness Checklist, notably with those items involving Probable Adaptation and Achievement. As with the summer administrations, correlations between these checklist items and summary scores for Verbal-Social Participation, Achievement Motivation and Independence were high and positive, those involving the summary score of Timidity was high and negative, while correlations involving Aggression were moderate to low and negative, indicating a somewhat lessened tendency for aggression to be seen as seriously hindering future adaptation and achievement.

As mentioned above, four items included in the retest administration of the Readiness Checklist (Appearance, Motor Coordination, Speech and Engagement) were originally a part of the Naturalistic Observations (see Chapter V), and were rated there on the basis of behavior during the time sample being observed, rather than on general, long-term observations of the child. During the summer, approximately sixteen such ratings were made on each child, and an average score for each of the four items was later computed.

In looking at correlations between these items (e.g. Appearance, Motor Coordination, Speech and Engagement) and summer administrations of the Behavior Inventory, it was noted that the magnitude of correlations varied considerably by variable, but they correlated equally well or poorly with teacher or observer administrations of the Behavior Inventory.

High ratings on these four variables correlated negatively with high ratings on Aggression and Timidity, positively with high ratings on Achievement Motivation, Verbal-Social Participation and Independence. The majority of correlations obtained were significant at better than the .02 level (see Table III-1). Correlations between the retest administration of the items including Appearance, Motor Coordination, Speech and Engagement (included in the Readiness Checklist and rated on the basis of general rather than immediate gehaviors) and the retest administration of the Behavior Inventory followed the trend described above. Here, however, correlations were strikingly higher than those between summer administrations of the Behavior Inventory and the four items listed above, then a part of the Naturalistic Observations, indicating either differential inter-rater perceptions or differences in ratings based upon long-term versus short-term observed characteristics.

Correlations between Summary Scores from the summer administration of the Binet Face Sheet by testers and the summer administration of the Readiness Checklist by teachers and observers were positive, but ranged widely in magnitude (see Table III-1). The Binet Summary Score for Achievement Motivation correlated highly with all summer ratings of Kindergarten Readiness, Adaptation and Achievement. While all correlations between summer Readiness Checklist items and summer Binet Summary Scores for Confidence in Ability and Activity Level were positive, only two were significant at better than the .05 level ($r = .240$ for Confidence in Ability vs. the Readiness Checklist item for Achievement as rated by teachers; $r = .353$ for Activity Level vs. Achievement by teachers).

All correlations between the retest administration of the Readiness Checklist and the Binet Face Sheet by teachers were positive and significant. High Achievement Motivation, Confidence in Ability and Activity Level, rated by kindergarten teachers on perceived long-term rather than upon situation-specific characteristics, relate to perceived Kindergarten Readiness, Achievement, Adaptation, neat appearance, coordinated motor behavior, frequency of speech and degree of engagement.

For the summer administration of the Readiness Checklist and the Binet Face Sheet, two factors may have operated to depress the magnitude of the correlations: (a) differences in inter-rater (teacher-tester) perceptions and (b) differences in ratings based upon long-term experience with the child in many situations versus short-term observation of his behavior in a testing situation. It is impossible to separate out effects of these two factors, as in no instance were long-term vs. short-term ratings on the same instrument made by the same individuals.

3. Summary

The Readiness Checklist provides an essential link between behavioral and cognitive areas; it coreelates highly and significantly with such cognitive variables as Stanford-Binet I.Q., Reading and Number Readiness, and the Preschool Inventory, and, as we have shown, it also correlates significantly with other behavioral measures. (see Table III-2)

The items from the original instrument which have given evidence of highest predictive value are: Kindergarten Readiness, Adaptation, and Achievement. Items which were later added (Appearance, Motor Coordination, Speech and Engagement) are also valuable, and it is recommended that future programs include these seven items.

TABLE III-1

Correlations of Readiness Checklist with Other Behavioral Measures[a]

SUMMER →

	Kindergarten Readiness Teacher r	Adaptation Teacher r	Achieve. Teacher r	Adapt. Observer r	Achieve. Observer r
A. Behavior Inventory Summary Scores					
I. Summer, Teacher I					
Aggression	-.619**	-.610**	-.400**	-.214**	-.173**
Verb-Soc Participation	.389**	.361**	.407**	.428**	.292**
Timidity	-.439**	-.455**	-.445**	-.367**	-.283**
Independence	.437**	.485**	.447**	.253**	.207**
Achievement Motivation	.582**	.574**	.564**	.283**	.338**
	N = 159-170 →			N = 121-125 →	
II. Summer, Teacher II					
Aggression	-.569**	-.547**	-.255**	-.252**	-.183
Verb-Soc Participation	.382**	.371**	.390**	.419**	.326**
Timidity	-.458**	-.477**	-.441**	-.443**	-.321**
Independence	.450**	.547**	.491**	.316**	.311**
Achievement Motivation	.608**	.636**	.498**	.371**	.384**
	N = 124-131 -->			N = 88-95 -->	
III. Summer, Observer I					
Aggression	-.401**	-.279**	-.124	-.262**	-.166
Verb-Soc Participation	.218**	.220**	.336**	.392**	.410**
Timidity	-.335**	-.308**	-.247**	-.585**	-.414**
Independence	.196**	.179*	.101	.273**	.206**
Achievement Motivation	.374**	.364**	.354**	.534**	.524**
	N = 160-169 -->			N = 123-126 -->	
IV. Retest, Teacher					
Aggression	-.217*	-.165	-.208	.128	.150
Verb-Soc Participation	.393**	.399**	.460**	.330**	.259*
Timidity	-.270**	-.254**	-.298**	-.374**	-.179
Independence	.480**	.541**	.444**	.311**	.293**
Achievement Motivation	.352**	.409**	.427**	.260*	.185
	N = 81-84 -->			N = 64-66 -->	
B. Binet Face Sheet Summary Scores					
I. Summer, Tester					
Achievement Motivation	.427**	.394**	.487**	.269**	.374**
Confidence in Ability	.216**	.201**	.240*	.072	.142
Activity Level	.138	.192*	.353**	.085	.171
	N =122-146-->			N = 99-109-->	
II. Retest, Teacher					
Achievement Motivation	.371**	.376**	.414**	.333**	.290**
Confidence in Ability	.314**	.270**	.405**	.334**	.338**
Activity Level	.277**	.223*	.356**	.271*	.228*
	N = 81-85 -->			N = 63-67 -->	

[a] Product-moment correlations based upon working-class children from Centers A and B.

* = significant at .05

** = significant at .02

TABLE III-1

SUMMER ---

	Appearance	Motor	Speech	Engagement
A. Behavior Inventory Summary Scores	r	r	r	r
I. Summer, Teacher I				
Aggression	-.342**	-.410**	-.177	-.271**
Verb-Soc Participation	.266**	.209	.418**	.187
Timidity	-.272**	-.210	-.467**	-.176
Independence	.037	.164	.285**	.312**
Achievement Motivation	.281**	.313**	.493**	.306**
	N = 70-74 --→			
II. Summer, Teacher II				
Agression	-.282**	-.414**	-.262*	-.282**
Verb-Soc Participation	.271**	.269**	.434**	.189
Timidity	-.252*	-.160	-.411**	-.091
Independence	.115	.259*	.322**	.361**
Achievement Motivation	.231*	.349*	.424**	.355**
	N = 70-74 --→			
III. Summer, Obs. I				
Aggression	-.218	-.485**	-.202	-.254*
Verb-Soc Participation	.088	.187	.300**	.134
Timidity	-.169	-.181	-.383**	-.139
Independence	.188	.266*	.187	.269*
Achievement Motivation	.328**	.345**	.350**	.375**
	N = 69-73 --→			
IV. Retest, Teach.				
Aggression	-.385**	-.512	-.228	-.217
Verb-Soc Particip.	.249*	.304**	.637**	.371**
Timidity	-.320**	-.232*	-.516**	-.076
Independence	.262*	.413**	.556**	.557**
Achievement Motivation	-.331**	-.416**	-.483**	-.558**
	N = 69-72--→			
B. Binet Face Sheet Summary Scores				
I. Summer, Tester				
Achievement Motivation	.218	.520**	.601**	.535**
Confidence in Ability	-.069	.168	.299**	.222
Activity Level	-.027	.188	.331**	.165
	N = 99-109 --→			
II. Retest, Teacher				
Achievement Motivation	.109	.476**	.486**	.233*
Confidence in Ability	.389**	.358**	.529**	.261*
Activity Level	.229	.145	.459**	.155
	N = 63-71 --→			

* = significant at .05
** = significant at .02

TABLE III-1

FALL RETEST -->

	Kinder-garten Readiness	Achieve-ment	Adapta-tion	Appear-ance	Motor Coordi-nation	Speech	Engage-ment
A. Behavior Inv. Summary Scores	r	r	r	r	r	r	r
I. Summer, Teach. I							
Agression	-.298**	-.240*	-.357**	-.162	-.221*	-.055	-.194
Verb-Soc Partic.	.278**	.265**	.398**	.231*	.356**	.443**	.316**
Timidity	-.185	-.273**	-.381**	-.169	-.244**	-.403**	-.289**
Independence	.348**	.349**	.345**	.073	.341**	.196	.314**
Achieve. Motiv.	.376**	.477**	.507**	.335**	.323**	.283**	.369**
	N = 87-90 -->						
II. Summer, Teach.II							
Aggression	-.278**	-.313**	-.417**	-.159	-.300**	-.081	-.265**
Verb-Soc Partic.	.188	.266**	.381**	.271**	.378**	.489**	.298**
Timidity	-.132	-.238**	-.347**	-.058	-.250**	-.365**	-.213*
Independence	.362**	.364**	.399**	.197	.406**	.235*	.359**
Achieve. Motiv.	.414**	.480**	.515**	.307**	.389**	.258**	.369**
	N = 70-74 -->						
III. Summer, Obs. I							
Aggression	-.262**	-.140	-.273**	-.211*	-.212*	.131	-.161
Verb-Soc Partic.	.212*	.159	.309**	.153	.300**	.337**	.197
Timidity	-.121	-.120	-.310**	-.067	-.285**	-.283**	-.146
Independence	.146	.192	.176	.096	.163	-.001	.150
Achieve. Motiv.	.324**	.315**	.394**	.285**	.301**	.160	.194
	N = 86-88 -->						
IV. Retest, Teach.							
Aggression	-.248**	-.309**	-.429**	-.228*	-.166	.031	-.248**
Verb-Soc Particip.	.410**	.623**	.642**	.532**	.593**	.818**	.453**
Timidity	-.187	-.296**	-.565**	-.268**	-.287**	-.527**	-.407**
Independence	.645**	.759**	.712**	.395**	.538**	.535**	.649**
Achieve. Motiv.	.520*	.687**	.580**	.284**	.400**	.409**	.504**
	N = 84-86 -->						
B. Binet Face Sheet Summary Scores							
I. Summer, Tester							
Achieve. Motiv.	.432**	.454**	.530**	.290**	.464**	.301**	.447**
Confid. in Ability	.222*	.277**	.170	.165	.276**	.287**	.113
Activ. Level	.312**	.173	.208	.170	.251**	.135	.212*
	N = 79-86 -->						
II. Retest, Teach.							
Achieve. Motiv.	.395**	.439**	.614**	.371**	.505**	.326**	.343**
Confid. in Ability	.246*	.312**	.435**	.222*	.328**	.394**	.280**
Activity Level	.106	.258*	.315**	.305**	.243*	.501*	.249**
	N = 82-87 -->						

* = significant at .05
**= significant at .02

Items from the original checklist which have shown no predictive or correlative value or which have not been analyzed are (a) the two like-dislike items, and (b) the nine items indicating specific areas in which the child was perceived as not being prepared for kindergarten. Regarding the like-dislike items, preliminary analysis indicated that neither teachers nor observers are able or willing to indicate more than moderate dislike for a child. This tendency may have been influenced by the very philosophy of the Head Start Program, as well as by the personal attitudes regarding culturally disadvantaged children of the adults who freely elected to work with them. Our teachers have shown strong and commendable interest in the problems and issues raised in the program; they tend to see tasks as challenging. This perspective, which works in the best interests of the Head Start Program, disinclines teachers to say they dislike a child.

The nine items indicating specific areas in which the child was perceived as not being prepared for kindergarten were omitted from this analysis. The overly general nature of the items, as well as the high and significant correlations obtained between other items on the Readiness Checklist and Behavior Inventory summary scores rendered these items somewhat superfluous. It was felt that similar but more sensitive measures of unreadiness for kindergarten could be obtained by already existing reliable instruments.

D. Stanford-Binet Face Sheet

The Face Sheet of the Stanford-Binet Intelligence Scale, Form L-M, consists of 13 items and one overall rating of conditions affecting test performance. Each of the thirteen sets of behavior during the test is rated on a nine-point scale, indicating high or low position on the continuum.

This instrument was, during the summer Head Start Program, administered to testers at the conclusion of the intelligence test, and was rated on the basis of behavior occurring during the testing situation. During the retest program, it was administered to teachers together with other behavioral ratings, and her # each child was rated on the basis of overall behavioral tendencies observed during the school term.

During the summer program, children from Centers A and B were rated on this instrument; during the retest program, children from Center A only received these ratings.

1. Formation of Summary Scores for the Binet Face Sheet

Two factor analyses, one including all observations for the summer administration to testers (N = 190-211) and the other including the retest administration to teachers (N = 268-274 - this includes both Head Start and non-Head Start children in Center A) were conducted using these thirteen items, excluding the overall rating, for the purpose of determining summary areas for computing subscores and reducing the number of items. Three factors were extracted using a Principal Component Analysis. In each factor, items with the highest loading were selected with the additional criteria that loadings for teachers and testers be

TABLE III - 2

CORRELATIONS OF READINESS CHECKLIST ITEMS WITH COGNITIVE MEASURES+

1. Readiness Chklist	Stanford Binet	Preschool Inventory Partial Set Sum. 1	Preschool Inventory Partial Set Sum. 2	Retest	Nat.% rank Ach.Test Spr.'66	Number Readiness	Reading Readiness
	r	r	r	r	r	r	r
I. Summer, Teacher							
Kinderg. Readiness	.328**	.608**	.749**	.555**	.528**	.528**	.545**
Adaptation	.226**	.516**	.618**	.486**	.575**	.548**	.555**
Achievement	.408**	.605**	.654**	.596**	.628**	.612**	.678**
II. Summer, Observer							
Adaptation	.344**	.400**	.580**	.504**	.500**	.473**	.451**
Achievement	.477**	.517**	.649**	.523**	.575**	.592**	.565**
	N = 117 - 165-->		N=45-61	N=68-88	N = 56-73-------->		
III. Retest, Teacher							
Kinderg. Readiness	.544**	.575**	.609**	.628	.475**	.507**	.491**
Adaptation	.618**	.673**	.602**	.670**	.588**	.606**	.604**
Achievement	.563**	.702**	.657**	.598**	.650**	.626**	.658**
Appearance	.345**	.356**	.196	.388 **	.551**	.445**	.507**
Motor Coordination	.515**	.544**	.449**	.461**	.571**	.493**	.562**
Speech	.474**	.405**	.345*	.353**	.431**	.308**	.408**
Engagement	.536**	.578**	.501**	.615**	.366**	.323**	.374**
	N = 81-87----->		N=28-29	N=89-90	N = 72-73--------->		

+Correlations here presented are based on workimg-class children from Centers A and B

* = significant at p = .05 or better
** = significant at p = .02 or better

similar. Summary Scores 1 and 2 each contain four items; Summary Score 3 contains three items. Two of the original thirteen items were not included in the summary scores; one of these, the rating of activity level (Normal activity level.... Hyperactive or depressed) was loaded on Factor 1 .723 for teachers' ratings and .497 for testers. By the criterion that this item is really tri-polar, not bipolar, and because it did not fit in with the remainder of the items in content, this item was not included in this summary score. The second item excluded (Reacts to failure realistically....withdrawing, hostile, denying) was loaded on Factor 1 .295 for teachers and .616 for testers, and on Factor 3 .780 for teachers and .258 for testers. Because this item did not match in factor structure for the two groups, it was not included.

Table III-3 lists the items included in each of the 3 Summary Scores. It is recommended for future studies of this nature that the Binet Face Sheet include only these 11 items, and that they be summed to form these Summary Scores. All of the analysis reported here for this instrument uses only these 3 Summary Scores.

2. Correlation of Binet Face Sheet Summary Scores with other Measures of Behavior

Summary scores from the Binet Face Sheet correlate moderately with three of the five item clusters from the Behavior Inventory. High Confidence, Activity Level and Achievement Motivation on the Binet Face Sheet correlate positively and, in the majority of cases, significantly at better than the .02 level, with high Verbal-Social Participation and Achievement Motivation on the Behavior Inventory, and negatively with high Timidity, both on summer compared to summer and retest compared to retest administrations of the two instruments, with correlations between retest administrations being generally higher than between summer administrations (see Table III-4). For both administrations, correlations between Binet summary scores and Summary Scores from the Behavior Inventory involving Aggression and Independence tend, with few exceptions, to be insignificant, though certain trends are indicated: high ratings on Aggression in some cases are negatively correlated with high Binet ratings on Achievement Motivation, and positively correlated with high Binet ratings on Confidence and Activity Level. High ratings on Independence tend to be positively correlated with high Binet ratings on Achievement Motivation, Confidence, and Activity Level.

Correlations between the Binet Summary Scores and items on the Readiness Checklist tend to be more substantial (see Table III-1). The Binet summary score for Achievement Motivation (summer administration to testers) correlated highly and significantly (p = .02 or better) with all summer ratings by both teachers and observers of Kindergarten Readiness, Adaptation and Achievement. While all correlations between Binet Summary Scores for Confidence in Ability and Activity Level and Readiness Checklist items were positive, only two were significant at better than the .05 level (r = .240 for Confidence in Ability vs. the Readiness Checklist item for Achievement as rated by teachers; r = .353 for Activity Level vs. Achievement by teachers).

All correlations between the retest administrations of the Binet Face Sheet and the Readiness Checklist by teachers were positive and significant. High Achievement Motivation, Confidence in Ability and Activity Level, here rated on perceived

TABLE III - 3

ITEMS INCLUDED IN SUMMARY SCORES FOR STANFORD BINET FACE SHEET RATINGS BASED ON BOTH TEACHERS AND TESTERS+

Summary Score 1 -- Achievement Motivation

Loading on Rotated Factor 1 for Tester	Loading on Rotated Factor 1 for Teacher	Item and Test Makers Designation
.759	.673	Challanged by hard task...Prefers only easy task (Problem Solving).
.719	.747	Persistent...Gives up or can't give up easily (Problem Solving)
.650	.700	Eager to continue...Seeks to terminate (ProbS)
.517	.816	Absorbed by task...Easily distracted (Attention)

Summary Score 2 -- Confidence in Ability

Loading on Rotated Factor 2 For Tester	Loading on Rotated Factor 2 for Teacher	Item
.718	.493	Comfortable with adults...Ill-at-ease (Emot.Indep
.713	.796	Assured...Anxious about success (Emotion. Indepen
.653	.631	Realistically self-confident...Distrusts own ability or over-confident (Emotional Independ.
.463	.644	Needs minimum commendation...needs constant praise and encouragement (Independence of Examiner Support)

Summary Score 3 -- Activity Level

Loading on Rotated Factor 3 for Testers	Loading on Rotated Factor 3 for Teacher	Item
.833	.871	Initiates Activity...Waits to be told (Reactions during Test)
.716	.643	Socially confident...Shy, reserved, reticent (Emotional Independence)
.695	.598	Quick to respond...Urging needed Reactions during Test)

+ The factor analysis on which this table is based was a Principal Component Analysis of the 13 items for testers and for teachers separately. Varimax Rotation of the 3 factor solution is cited. For testers, N ranges from 190 to 211. For teachers, N ranges from 268 to 274.

long-term rather than upon situation-specific characteristics, relate to perceived Kindergarten Readiness, proposed Achievement and Adaptation, neat appearance, coordinated motor behavior, frequency of speech and degree of engagement.

As we have indicated in the section of the report dealing with the Readiness Checklist, for the summer administrations of both instruments, two factors may have operated to depress the magnitude of the correlations: (a) the perceptions of different raters and (b) differences in ratings based upon long-term experience with the child in many situations versus short-term observation of his behavior in a testing situation. It is impossible to separate out effects of these two factors, as in no instance were long-term vs. short-term (test specific) ratings on the same instrument made by the same individuals. There is also the possibility that testers rated these children on a baseline of ability rather than upon behavior patterns, as testers rated these children immediately following the intelligence test and could not help but be aware of ability demonstrated in the testing situation. As the product-moment correlation between testers' ratings of Achievement Motivation on the Binet and the I.Q. score was .626 (as opposed to a correlation of .497 for I.Q. versus retest ratings by teachers on Binet Achievement Motivation), there is some evidence for the tester's use of this rating scale for their own estimate of how well the child actually performed on the test, not how motivated he was to achieve.

3. Correlation of Binet Face Sheet Summary Scores with Cognitive Measures

Product-Moment correlations between summary score items on the Binet Face Sheet and such cognitive variables as I.Q., number correct on the Preschool Achievement Test, National Percentile Rank Achievement Test, Number Readiness and Reading Readiness were, with few exceptions high and significant at better than the .02 level (see Table III-5). High Achievement Motivation, Confidence in Ability, and Activity Level seem thus to be associated with high cognitive performance. The high magnitude of some of the correlations may, however, be due to I.Q. That is, Scores on the Preschool Achievement Test are highly correlated with I.Q.; the Binet Face Sheet Summary Score for Achievement Motivation is also highly (and perhaps artificially for the testers) related to IQ. Therefore, the correlation between Achievement Motivation and the Preschool Achievement Test may be spurious.

Perhaps the magnitude of the majority of these correlations has been affected by teacher or tester bias, or perhaps by relationships between certain cognitive measures. The fact, however, that other behavioral measures as the Readiness Checklist and the Behavior Inventory, as rated by theoretically less biased observers, show high correlations with these same cognitive measures lends weight to the behavioral-cognitive tie.

4. Stanford-Binet Face Sheet: Summary

These eleven items have been found useful in differentiating children along both behavioral and cognitive lines, and it is recommended that they be included in future studies. If the Binet is given, testers should fill these items in; they are also useful for general rating scales. The items themselves differ from the majority of items on the Behavior Inventory in their emphasis on behavior

TABLE III - 4

CORRELATIONS OF BINET FACE SHEET SUMMARY SCORES WITH BEHAVIOR INVENTORY SUMMARY SCORES

A. Behavior Inventory Summary Scores	Binet, Summer, Testers Adm.			Binet, Fall Ret. Teachers Adm.		
	Achievement Motivation	Confidence in Ability	Activity Level	Achieve. Motivat.	Confid. in Ability	Activity Level
I. Summer, Teacher I						
Aggression	-.150	.012	.113	-.235*	.054	.135
Verbal-Soc. Participation	.284**	.224**	.338**	.290**	.473**	.447**
Timidity	-.308**	-.300**	-.290**	-.265**	-.405**	-.409**
Independence	.144	.145	-.017	.192	.005	-.040
Achieve.-Motivation	.324**	.231**	.154	.199	.258**	.197
	N = 129-153--------------->			N = 81-87------------>		
II. Summer, Teacher II						
Aggression	-.211	.006	.127	.259**	-.024	.057
Verbal-Soc. Particip.	.397**	.252**	.290**	-.308**	-.546**	.513**
Timidity	-.348**	-.265**	-.262**	.232 *	.428**	-.425**
Independence	.260**	.129	.078	-.251**	-.162	.070
Achieve. Motivation	.395**	.252**	.176*	.215*	.253**	.179
	N = 110-127------------->			N = 82-87------------>		
III. Summer, Observer I						
Aggression	-.042	.075	.106	-.246*	-.002	.029
Verbal-Soc Particip.	.326**	.190*	.398**	.346**	.421**	.474**
Timidity	-.252**	-.271**	-.276**	-.273**	-.326**	-.345**
Independence	.124	.137	-.018	.159	.064	.069
Achieve. Motivation	.287**	.219**	.184*	.254**	.300**	.215*
	N = 133-155---------- ----->			N = 80-85------------->		
IV. Fall Retest, Teacher						
Aggression	-.168	.057	.025	-.178	-.157	.098
Verbal-Soc Particip.	.451**	.278**	.192	.436**	.430**	.515**
Timidity	-.288**	-.026	-.221*	-.347**	-.427**	-.389**
Independence	.441**	.221	.127	.421**	.252*	.138
Achieve, Motivation	.269**	.166	.118	.343**	.298**	.230*
	N = 75-83------------------>			N = 79-83------- - -->		

Product-moment correlations based on working-class children from Centers A and B.

For Correlations of Binet Summary Scores vs. the Readiness Checklist, see Table III-1

* = significant at p = .05 or better
** = significant at p = .02 or better

TABLE III - 5

CORRELATIONS OF BINET FACE SHEET SUMMARY SCORES WITH COGNITIVE MEASURES

		Preschool Inventory					
Binet Face Sheet Summary Scores+	Stanford Binet	Partial Set Sum. 1	Partial Set Sum. 2	Retest	Nat.%rank Ach. Test Spr. '66	Number Readiness	Reading Readiness
I. Summer, Tester							
Achievement Motivat.	.626**	.561**	.527**	.538**	.520**	.516**	.545**
Confidence in Ability	.424**	.326**	.230	.318**	.367**	.279*	.358**
Activity Level	.399**	.313**	.280*	.377**	.330**	.264*	.330**
	N = 118-158---		N=42-53	N=80-86	N=67-71-------------		
II. Retest, Teacher							
Achievement Motivat.	.497**	.489**	.632**	.629**	.593**	,572**	.621**
Confidence in Ability	.448**	.438**	.574**	.586**	.525**	.419**	.536**
Activity Level	.378**	.375**	.462**	.517**	.462**	.344**	.486**
	N=77-85----		N=26-28	N=83-87	N=68-71-------------		

+Correlations here presented are based on working-class children from Centers A and B.

* = significant at p = .05 or better
** = significant at p = .02 or better

during task situations, rather than on more general, social or interpersonal situations. They differ from items on the Readiness Checklist in their lessened emphasis on general school adjustment.

E. Preschool Inventory: Restrictive-Supportive Items

We have included the analysis of this part of the Preschool Inventory (PAT) in this section of the report as we felt that these items not only apply more directly to the personality and behavioral areas than they do to the cognitive, but that they also involve direct responses by the child rather than ratings of the child by others. We felt, moreover, that there should be a relationship between actual behavior and perception of authority figures as restricting or supporting the self.

Eight of the questions on Dr. Bettye Caldwell's Preschool Inventory, which was designed specifically for the Head Start program, involve the perceived function of various authority figures in the home and community. The questions are:

What does a doctor do?
What does a dentist do?
What does a father do?
What does a mother do?
What does a policeman do?
What does a teacher do?
What does a nurse do?
What does a soldier do?

The answers given by the child to these questions were originally to be scored on a two-point restrictive-supportive scale (see Caldwell, 1965). We found that the original instructions for scoring these items were not adequate for our needs, as not only did we find what we considered to be a range in degree of restrictive or supportive responses, but we found neutral responses as well. We felt, for example, that a response of "She whips you" to the question of "What does a teacher do?" was a decidedly more restrictive one than the response "She makes us sit down." A response that a father "Watches T.V." or "Drinks beer" was felt to be more neutral than either supportive or restrictive. We therefore constructed a five-point restrictive-supportive scale, with the ends of the continuum representing extreme degrees of restrictiveness or supportiveness, the second and fourth points representing moderate degrees of orientation, and the middle point representing neutrality.

1. Intercorrelations of Restrictive-Supportive Items

Although extensive analysis has not yet been completed with this set of items, some correlational material is availabe. Product-moment correlations between items on the initial summer administration of this test by teachers do not, on the whole, reach the .05 level of significance. Some correlations obtained, however, were significant (see Table III-6); A few significant negative correlations were found between number of words and restrictive-supportive scores, indicating that children responding positively tended to speak at greater length than did those who responded negatively. Significant correlations were found between the eight restrictive-supportive scores for each child, indicating a certain consistency of response orientation along this dimension.

TABLE III - 6

SIGNIFICANT CORRELATIONS BETWEEN SELECTED RESTRICTIVE-SUPPORTIVE ITEMS+

Restrictive-Supportive Score	What does a doctor do? Restrict.-Supp.	What does a policeman do? Restrict.-Supp.
Policeman	.175*	
Dentist	.215*	.295**
Teacher	.247**	.258**
Father	.242**	
Nurse	.286**	
Mother	.217*	
Soldier	.196*	

+These correlations are based on working class children from Centers A and B. The N ranges from 111-154.
A complete intercorrelation matrix has not been included, as additional item comparisons differ little from these presented.

* = significant at p = .05 or better
** = significant at p = .02 or better

2. Correlation of Restrictive-Supportive Items with Behavior Inventory Summary Score

Product-moment correlations were obtained between restrictive-supportive items from the initial summer administration by teachers and Summary Scores on the Behavior Inventory. Significant correlations are scattered, but they are consistent and do indicate certain trends (see Table III-7). As they occur with equal frequency and magnitude between the initial administration of these Preschool Inventory items and teacher, observer, and retest administrations of the Behavior Inventory, we shall summarize them.

Restrictive-supportive scores tended to correlate negatively and significantly with Behavior Inventory Summary Scores of Verbal-Social Participation, Independence, and Achievement Motivation, and positively and significantly with Timidity and, in a few cases, Aggression. High Aggression and Timidity were correlated positively with perceptions of figures as Restrictive; high Independence, Verbal-Social Participation, and Achievement Motivation were correlated with perceptions of figures as Supportive. Perception of authority figures, then, is related to actual observed behavior in much the same way as is cognitive performance.

3. Restrictive-Supportive Items: Summary

Extensive analysis of the eight items on the Original Preschool Inventory involving perception of authority figures as restrictive or supportive has not been completed. We have found significant intercorrelations both between the restrictive-supportive items themselves and between the restrictive-supportive items and Summary Scores on the Behavior Inventory. Analysis involving the formation of summary scores for the restrictive-supportive items is planned, and results of this analysis will be available at a future date.

F. Behavior Inventory

The Behavior Inventory, originally a fifty-item instrument, was designed to measure certain behavioral and emotional tendencies ranging from verbal participation, social interaction and aggression to general dispositional states. Each child was rated for each item on a 7-point scale; numerically low ratings indicate similarity to or possession of the attribute in question, numerically high ratings indicate dissimilarity. The original instrument was administered four times, once to teachers and once to observers at the onset of the Head Start Program, and again to both teachers and observers during the eighth week of the program. The teachers' initial administration of the instrument included children from Centers A, B and C; the initial observers' administration included children from Centers A and B. The second teachers' administration included from Centers A and B, and the second observers' administration included only children from Center A. During the retest program, a condensed version of the instrument was administered to teachers in Center A.

The instrument was also administered to teachers, with ratings based on a 4-point scale, at the onset and at the conclusion of the program. The initial

TABLE III - 7

SIGNIFICANT CORRELATIONS BETWEEN SELECTED RESTRICTIVE-SUPPORTIVE ITEMS AND BEHAVIOR INVENTORY SUMMARY SCORES+

	What does a doctor do? Restrictive	What does a policeman do? Restrictive	What does a teacher do? Restrictive
Aggression	.210*		.244*
Verbal-Social Participation		-.422**	
Timidity	.230*	..204*	
Independence	-.255**		
Achievement Motivation	-.295**	-.290**	

+Correlations presented are based on working class children from Centers A and B. N ranges from 91-117.

* = significant at p = .05 or better
** = significant at p = .02 or better

administration here included children from Centers A, B, C and D, and the second administration included children from Centers A and B.

1. Formation of Summary Scores for Behavior Inventory

On the basis of preliminary factor analyses, twenty-three items were chosen for followup testing during the autumn following the Head Start summer. The major criterion for including an item in the retest was its high loading on one of the rotated factors. A few items were included to sample some general positive behavior characteristics even though these items had not loaded particularly highly on any factor.

A more complete factor analysis including all observations (N = 769) for teachers and observers summer testings was conducted using only these twenty-three selected items, for the purpose of determining summary areas to compute subscores and reduce the number of items. Six factors were extracted using a Principle Component Analysis. For the first five factors the four items with the highest loadings were selected differently from all others in the Behavior Inventory.

These suggested summary scores are not factor scores in the true sense because items included were not weighted by their loadings on the factor (although the item which was loaded negatively on the third factor was reversed in scoring). Table III-8 lists the items included in each of the 5 summary scores. It is recommended in the future that the Behavior Inventory include only these 20 items, and that they be summed to form summary scores. All of the analysis reported here for the Behavior Inventory uses only these five summary scores.

2. Comparisons of Results on Behavior Inventory Using Four and Seven Point Scales

The original Behavior Inventory as sent out by the Office of Economic Opportunity required that items be rated on a four point scale (1. Very much like; 2. Somewhat like; 3. Very little like; 4. Not at all like). Because the research staff felt that this scale did not allow for sufficient discrimination, particularly for someone who was like the statement, but neither as much as "very much like" or as little as "somewhat like," a seven point rating scale was applied to every child who was rated (1. Exactly like; 2. Very much like; 3. Quite a bit like; 4. Pretty much like; 5. Somewhat like; 6. Very little like; 7. Not at all like). In addition, 136 protocols were filled out by teachers using the 4-point scale.

The correlations between the application of the 4-point and the 7-point scales to the same child for the same scale ranged from .70 to .94 (N ranged from 132 to 136), for the 50 scales used in the total Behavior Inventory.

As a more stringent test of whether the additional variance and discrimination provided by the 7-point scale provided additional useful variance and subject discrimination, Two Principal Componant Factor Analyses were made of the 50 items using only the 136 teachers who had answered both forms of the questionnaire; one factor analysis was done using all the 4-point scales, another using all the 7-point scales. The rotated factor solutions were very similar, in spite of the fact that the analyses were performed independently and there was no attempt to

TABLE III - 8

ITEMS INCLUDED IN SUMMARY SCORES FOR THE BEHAVIOR INVENTORY AND THEIR FACTOR LOADINGS+

Summary Score 1 -- Aggression

Loading on Rotated Factor 1	Item
.816	Responds to frustration or disappointment by becoming aggressive or enraged
.781	Emotional response is customarily very strong: over-responds to usual classroom problems, frustrations.
.780	Is often quarrelsome with classmates for minor reasons.
.752	Little respect for the rights of other children (won't wait turn, etc.)

Summary Score 2 -- Verbal-Social Participation

Loading on Rotated Factor 2	
.765	Talks eagerly to adults about his own experiences and what he thinks.
.764	Likes to talk or socialize with teacher.
.735	Is eager to inform other children of the experiences he has had.
.691	Asks many questions for information about things, persons, etc.

Summary Score 3 -- Timidity

Loading on Rotated Factor 3	
.754	Is constricted, inhibited, timid: needs urging to engage in activities.
-.656	Is usually carefree: rarely frightened or apprehensive.
.656	Often keeps aloof from others because he is uninterested, suspicious, bashf
.607	Often will not engage in activities unless encouraged.

(continued on next page)

Summary Score 4 -- Independence

Loading on Rotated Factor 4	
.777	Tries to figure out things for himself before asking adults or children for help.
.719	Does not need attention or approval from adults to sustain him in his work.
.598	Appears to trust in his own abilities.
.564	Goes about his activities with a minimum of assistance from others.

Summary Score 5 -- Achievement Motivation

Loading on Rotated Factor 5	
.765	Discontinues activity after exerting a minimum of effort.
.652	When faced with a difficult task, either does not attempt it or gives up quickly.
.643	Seems disinterested in the general quality of his performance.
.406	Is lethargic or apathetic; has little energy or drive.++

+ The factor analysis on which this table is based was a Principle Component Analysis of 23 items with Verimax Rotation of 6 factors. The sixth factor included only one item with high loading and was eliminated. The number of observations on which the correlation matrix and factor analysis were based was 769.

++ This item had a higher loading on Factor 3, but it was the 5th highest item on that factor, and only four items were to be included. It is included in Summary Score 5 because it was the fourth highest item loading on this factor, because it matched in content the other items included in the score, and for the sake of balancing the number of items per score.

Items Included in Retest but not in Summary Scores:

1. Insists on maintaining his rights (5th item on Factor 1).
2. Requires the company of other children (alone on Factor 6).
3. Demonstrates imaginativeness and creativity in the use of toys, etc. (5th item on Factor 4).

rotate them to a common structure. As would be expected, the first two factors matched each other more closely than succeeding ones. There was no consistent tendency for the communalities to be higher than the 7-point scale, and the factors had similar eigen values for the two analyses. The items as rated on the seven point scale were used in all reported analysis because the most extensive data had been collected using this item format. Although it is impossible to determine what results would have been obtained if the 4-point scale had been used, it is likely that the results would have been highly similar to those reported here.

3. Inter-rater Reliability

Inter-rater reliability (teachers vs. observers) on the Behavior Inventory was high to moderate, depending upon the specific variables involved with correlations being depressed to some extent by teachers in Center B.

In considering the entire sample of working class children available for Centers A and B, it was found that summary scores for Aggression, Verbal-Social Participation and Timidity showed highest Product-moment correlations ($p = .02$ or better) between teachers' and observers' initial ratings (see Table III-9), with lower correlations for summary scores for Independence and Achievement Motivation. It is evident from the above that some item clusters are more reliable in this respect than others, the less reliable clusters probably reflecting a certain ambiguity in the wording of the items themselves and differential perception of what constitutes "independent" or "achieving" behaviors.

In a previous analysis, it was noted that wide differences existed in teacher-observer correlations between Centers A and B. As four of the observers in Center A had also observed in Center B, it was decided to obtain new correlational material for the Center separately based upon observations made only by observers who had observed in both centers and the teachers (who were different for the two centers).

For Center A, these teacher-observer correlations were high, with item-cluster trends similar to those noted for correlations based upon the entire sample. For Center B, however, correlations were, with one exception, low, and no trend differentiating reliability by item cluster appeared.

In general, it appears that teachers in Center B have rated their children as showing less aggression, more verbal-social participation, less timidity, more independence, and more achievement motivation than have observers in that center (see Table III-10), while those same observers, who also observed in Center A, aggree in their ratings with teachers in Center A. A discussion of teacher differences in specific areas of experience and background is presented in Chapter IV of this report.

It would appear, then, with the exception of the group of teachers in Center B, whose ratings depressed correlations based upon the entire sample, that this instrument is reliable and can with confidence be used in future studies.

TABLE III - 9

BEHAVIOR INVENTORY: INTER-RATER RELIABILITY

A. Teachers vs. Observers, 1st Administration, Centers A and B Combined+

Summary Score	r	N
Aggression	.557**	173
Verbal-Social Participation	.568**	177
Lethargy	.564**	176
Independence	.328**	177
Achievement	.349**	176

+Product-moment correlations based on working-class children only - all observers

B. Teachers vs. Observers, 1st Administration Centers A and B Compared+

Summary Score	Center A		Center B	
	r	N	r	N
Aggression	.747**	47	.197	57
Verbal-Social Participation	.676**	48	.408**	59
Lethargy	.669**	48	.467**	59
Independence	.390**	48	.372**	59
Achievement	.404**	47	.160	60

+Product-moment correlations here are based on only working-class children observed by those observers who observed in both Centers A and B

** = significant at p = .02 or better

TABLE III - 10

BEHAVIOR INVENTORY SUMMARY SCORES: DISTRIBUTION OF RATINGS BY TEACHERS AND OBSERVERS IN CENTER B+

	Percentage of children receiving each Summary Score							
Rater	1.0-1.6	1.7-2.4	2.5-3.2	3.3-4.0	4.1-4.8	4.9-5.6	5.7-6.4	6.5-7.0
Summary Score 1 -- Aggression (high score equals less aggressive)								
Teacher	4.0	2.7	0.0	5.3	5.3	14.7	17.3	50.7
Observer	0.0	0.0	2.8	2.8	11.3	23.9	38.0	21.1
Summary Score 2 -- Verbal-Social Participation (high score = less verbal)								
Teacher	1.3	5.3	19.7	6.6	27.6	13.2	14.5	11.8
Observer	0.0	6.9	12.5	13.9	5.6	19.4	27.8	13.9
Summary Score 3 -- Timidity (high score = less timid)								
Teacher	0.0	2.6	10.5	5.3	23.7	22.4	21.1	14.5
Observer	1.4	2.8	8.3	6.9	13.9	30.6	22.2	13.9
Summary Score 4 -- Independence (high score = less Independence)								
Teacher	1.3	3.9	27.3	19.5	23.4	13.0	7.8	3.9
Observer	0.0	5.6	25.4	33.8	19.7	11.3	4.2	0.0

(continued on next page)

Summary Score 5 -- Achievement Motivation (high score = more achieving)

Rater	1.0-1.6	1.7-2.4	2.5-3.2	3.3-4.0	4.1-4.8	4.9-5.6	5.7-6.4	6.5-7.0
Teacher	0.0	0.0	3.9	7.8	3.9	19.5	40.3	24.7
Observer	0.0	0.0	2.8	5.6	16.7	29.2	38.9	6.9

+ The N of this sample ranges from 71 to 77 and is based upon working class children only.

From these distributions it can be seen that teachers in Center B rated their children as exhibiting less aggression, more verbal-social participation, less timidity, more independence and more achievement motivation than did observers.

Behavior Inventory: Changes in Summary Scores Over Time

At this point in our evaluation of the data, construction of change scores for each child for each of the five summary scores on the Behavior Inventory has not yet been completed. We must, then, look at available mean scores and their changes over time. Table III-11 presents means for each of the five summary scores for all teacher administrations of the instrument. Little change in the direction of more "socially acceptable" behavior patterns can be seen between the two summer administrations. Evidence of some change, though slight, is present, and occurs in the expected direction. Between the initial summer administration and the fall retest, however, evidence of change is clear-cut, with an average change of 0.80 along the 7-point scale, in the direction of more acceptable behavior.

This group of children, then, are perceived in the retest as being less aggressive, more verbal, less timid, more independent, and more achievement motivated than they were at the onset of the Head Start program. There is always the possibility that part or all of this perceived change may be due to rater bias; the differences between the initial and final summer administrations may be due to teachers' unconscious perception of change as unavoidable, that participation in Head Start must benefit a child. This factor, however, should not have entered into retest ratings, as officially these teachers did not know which children had had Head Start experience and which children had not. As the retest program was not observed, one has no way of knowing whether the summer-retest differences are due to teachers' bias. It is also possible that the less aggressive, more social, more independent children were the ones who stayed in the school system at Location A and therefore were retested. The differences between Head Start and non-Head Start children (presented in Chapter VII) substantiate the fact of some actual change in the child, rather than discrepancies in rater perceptions.

5. Behavior Inventory: Correlation with Cognitive Measures

Product-moment correlations between summer and retest administrations of the Behavior Inventory and such cognitive variables as Stanford-Binet I.Q., total correct on the Preschool Inventory (PAT), National Percentile Rank Achievement Test, Number Readiness and Reading Readiness are, with few exceptions, significant in the desired direction (see Table III-12). For all administrations of the Behavior Inventory, Summary Scores for Aggression and Timidity are negatively related to high cognitive performance, and Verbal-Social Participation, Independence, and Achievement Motivation are positively related. This instrument then, which is only minimally oriented to evaluation of task-related behavior, provides a more than moderate indication of cognitive performance; it is related to behaviors which are integrally related to academic success.

6. Behavior Inventory: Summary

This instrument has proved to have significant value in the Head Start research program, and it is highly recommended that it be included in future studies. It has provided a reliable basis for measurement of behavioral tendencies, and relates both to other behavior instruments and to cognitive measures as well. The twenty-item form of the instrument is easy to administer, and formulating summary scores for each of the five behavioral areas represented (see Formation of Summary Scores) greatly facilitates analysis.

TABLE III - 11

BEHAVIOR INVENTORY: CHANGES IN SUMMARY SCORES OVER TIME+

Summary Score	Teachers Time 1 Mean	N	Teachers Time 2 Mean	N	Teachers Retest Mean	N	Movement Over Time
Aggression	4.89	119	5.02	120	5.85	85	Less aggressive
Verbal-Social Participation	4.58	120	4.28	120	3.90	86	More Verbal
Timidity	4.64	119	4.74	120	5.43	86	Less Timid
Independence	4.10	120	4.12	120	3.59	86	More Independent
Achievement Motivation	4.94	118	5.15	120	5.74	85	More Achieving

+Means here presented are based on working-class children from Center A. Low ratings on summary scores for Aggression, Verbal-Social Participation, Timidity and Independence indicate similarity to or possession of the attribute in question. A low rating on the summary score for Achievement Motivation indicates an absence of achieving behavior.

TABLE III - 12

CORRELATIONS BETWEEN BEHAVIOR INVENTORY SUMMARY SCORES AND COGNITIVE MEASURES+

Behavior Inventory Summary Scores	Stanford Binet IQ	Preschool Inventory Partial Set Sum. 1	Preschool Inventory Partial Set Sum. 2	Retest	Nat % rank Ach. test Spr '66	Number Readiness	Reading Readiness
I. Summer, Teacher I							
Aggression	-.057	-.266**	-.344**	-.197	-.353**	-.356**	-.345**
Verb-Soc Participation	.282**	.340**	.432**	.360**	.274**	.208	.298**
Timidity	-.295**	-.386**	-.449**	-.409**	-.372**	-.287**	-.347**
Independence	.212**	.397**	.338**	.337**	.341**	.354**	.300**
Achieve.-Motivation	.171*	.414**	.407**	.518**	.468**	.385**	.493**
	N = 159-178----		N=65-67	N=88-90	N = 71-73--	------	
II. Summer, Teacher II							
Aggression	-.147	-.320**	-.441**	-.239*	-.353**	-.381**	-.339**
Verb-Soc Participation	.351**	.392**	.389**	.374**	.351**	.280**	.358**
Timidity	-.280**	-.386**	-.472**	-.385**	-.331**	-.293**	-.311**
Independence	.288**	.416**	.356**	.400**	.378**	.323**	.323**
Achieve.-Motivation	.282**	.463**	.469**	.523**	.476**	.434**	.451**
	N=126-138-----		N=58-59	N=89-90	N=72-73-------------		
III. Summer, Observer I							
Aggression	.014	-.206**	-.315**	-.273**	-.235*	-.244*	-.193
Verb-Soc Participation	.341**	.282**	.369**	.305**	.302**	.245*	.323**
Timidity	-.212**	-.328**	-.449**	-.353**	-.271**	-.206	-.240 *
Independence	.084	.174*	.214	.219*	.229*	.114	.131
Achieve.-Motivation	.168*	.310**	.392 **	.407**	.360**	.355**	.332**
	N=163-180-----		N=66-67	N=87-88	N=72---	----------	
IV. Retest, Teacher							
Aggression	-.282**	-.251*	-.173	-.128	-.226	-.276**	-.238 *
Verb-Soc Participation	.500**	.501**	.431**	.464**	.473**	.370**	.469**
Timidity	-.358**	-.353**	-.228	-.387**	-.296**	-.292**	-.250*
Independence	.498**	.573**	.547**	.456**	.526**	.532**	.527**
Achieve.-Motivation	.440**	.458**	.337	.425**	.470**	.535**	.460**
	N=77-83------		N=28-29	N=85-86	N=69-71------------		

+Correlations here presented are based upon workign-class children from Centers A &

* = significant at p = .05 or better
** = significant at p = .02 or better

CHAPTER IV

A DISCUSSION OF THE TEACHER AS A DETERMINANT IN THE HEADSTART PROGRAM

Abstract: Teacher behavior received little attention as an independent variable during the summer of 1965. Little correlation was found in a simple analysis of the relation of a child's kindergarten readiness or success in kindergarten with the summer Head Start teacher's educational level, experience with disadvantaged children, experience with pre-school children, or total number of years of teaching experience. However, a relationship was found to exist between these background variables and the extent to which teachers agreed with observers and the success with which they were able to predict a child's achievement in kindergarten. In general, previous experience -- especially with disadvantaged children and/or pre-school children -- increases the teacher's ability to predict accurately a child's adjustment in kindergarten. Such experience also tends to raise the agreement of the teacher with an observer.

Information from naturalistic observations indicated the importance of the teacher's actual performance in the classroom as a determinant of the child's behavior. Therefore, it is suggested that in the future attention should be focussed on the teacher and her performance in the classroom, and upon those variables in her background which affect that performance and her ability to evaluate children.

A study of the cognitive-stimulating aspects of the teacher's verbal behavior is now in progress.

A. Description of the Two Centers Involved in Study

1. General Characteristics and Physical Environment

There were important physical differences between the two centers followed in our study. Center A was well equipped and fully staffed -- in many ways a model physical setting. Classes were held in a modern public school which had a full range of Nursery-Kindergarten equipment, including a doll-corner, dress-up clothes for role playing, and a variety of cognitive games and art material. The outdoor playground had a variety of climbing equipment available as well as swings and slides, and even a "giraffe". Each class of fifteen was staffed by a teacher, and an assistant teacher and two or three volunteers.

In contrast Center B was housed in a church annex which had small rooms and a minimum of equipment. The playground consisted of a grassy lot with two trees for climbing but no swings or slides. In Center B there were two teachers for each group of thirteen with an occasional teen-age volunteer.

This chapter was prepared by Carla Berry

The overall programs for the two centers appeared to be similar: they both included free-play, group activities, eating, out-door play and art work. Center B, however, tended to have more group activity than Center A, and free play was less supervised. These contrasts reflect differences between the centers in resources as well as differences in goals.

An observer report characterized the aims of the Center A as "training for diffuse cognitive and social skills, 'creativity' and confident self-expression." These goals were implemented by the use of a carefully mediated environment which imparted a sense of freedom by having the structure implicit. This environment depended on flexibility provided by a high ratio of adult help (ready to step in to trouble shoot or to provide limits on an activity which was getting out of hand) and on play equipment which stimulated children to organize their own activities. As long as behavior was channeled within the broad confines, the children were not impinged on by rules and appeared to feel they were doing what they wanted. This type of programming maximized expressiveness and social interaction.

The aims of Center B were characterized by an observer as "training for school skills, obedience and, indirectly, social experience." The skills were related to specific information rather than to the enlargement of experience stressed in Center A. There also was more effort to check and evaluate the accomplishment of the children in mastering the skills. The classroom was more explicitly structured than in Center A. There were many definite rules which were strictly enforced and much less flexibility in the social groupings. A good portion of the day was spent in formal instruction on school skills and in organized games which involved the whole group. When the organized activities were not in progress there was a change to a higher degree of freedom than that at Center A.

A description of the "free play" illustrates some of these differences. Free Play is, of course, a misnomer. At Center A this term generally referred to "free selection" amongst carefully pre-chosen alternatives. In Center B "Free play" was structured by the children rather than by the available activities or the teachers. Often the teachers did not involve themselves at all but remained at their desks. Since there was little equipment to help the children structure their play, there was more physical activity and "horsing around" than at Center A. This was usually tolerated until rules were trespassed. In Center A all activity was perceived as relevant to generalized goals; involvement and informal teaching were present in all the child's experience, including "free play". In Center B the goals were more specific and achieved through more formal instruction.

2. The Staff

There were ten head teachers in Center A and four in Center B. In Center B, each teacher carried a morning and an afternoon session. Our information about these teachers includes: 1) a Pre-School Teacher Questionnaire; 2) a Workers Attitude Scale; 3) an observer's report on the classroom procedures; 4) an interview with the teacher; and 5) incidental observations of the teacher's behavior included in the naturalistic observations on the children.

A member of the staff is involved in differentiating teaching styles and describing the types of contact that teachers initiate with children. Categories are being developed to code this type of behavior. These will be described in subsequent reports.

a) Characteristics of the Teachers

The teachers in the two centers may be contrasted as groups. The average age of Center A teachers was higher (36 years) than that of teachers in Center B (27 years). The Center A teachers had a mean of 9.5 years of teaching experience, compared to 3.5 for Center B. Half of the four Center B teachers had only one year of teaching experience, whereas only one out of ten Center A teachers had only one year. Two of the four Center B teachers had had no previous experience with Nursery or Kindergarten children, whereas three out of ten Center A teachers had not had such experience. Three of the four Center B teachers had experience with disadvantaged children, while four of eight (two did not answer) in Center A had prior experience with disadvantaged children.

In summary, the Center B teachers were younger and had considerably less teaching experience especially with young children. On the other hand, they had a higher proportion of teachers who had worked with disadvantaged children, though the N is very small.

According to observers' reports the general atmosphere of the groups also varied. Center A teachers focussed on general characteristics and developmental needs of young children and saw their goals as aiding the children in the diffuse area of social facility. This was done through enrichment of general experience which may have included school skills. In Center B the focus was on the nature of disadvantaged children, who were perceived as having information deficits which might be corrected so that the children would be competitive with the better advantaged. Hence, the goals were more specific and were tailored to school demands manners, speech, and hygiene were also given high priority. These differences in goals and concern were useful for understanding the differences in structure and teaching in the two groups.

B. Analysis of the Data-1965

At this time it was not possible to pursue the qualitative differences described by the observers. The information available on all teachers was essentially demographic. Comparisons of teachers were made using information from the original Head Start Teacher Questionnaires. Simple correlations of demographic variables such as 1. the level of the teacher's educational training, 2. the number of years the teacher had worked with disadvantaged children, 3. the number of years experience the teacher had with pre-school children, and 4. the total number of years teaching experience, show no correlation with the Head Start child's readiness for kindergarten. A comparison of the teacher as an evaluator and predictor was also made with the observers. The results of this comparison appear in Chapter III. It was found that teachers in Center A tended to have a better correlation with observers on the Behavior Inventory than did teachers in Center B.

A second analysis using the demographic variables was done to see whether teacher differences existed which did not appear in the simple correlation matrix. High and Low groups were separated on each of the four variables, and the product-moment correlations on predictive and criteria variables for the sub-groups were compared. (see Table IV-1). This analysis indicates that differences do exist between the High and Low sub-groups. Caution is necessary in interpreting these figures as they are based on a very small teacher population. Therefore, the figures in Table IV-1 are presented as indicators for future investigations rather than as conclusive statistical evidence.

The two questions which were investigated were:

1) Do differences in the level of training and experience affect the correlation between observers' and teachers' evaluation of the same children?

2) Do differences in training and experience affect the success with which teachers are able to predict a child's achievement in kindergarten?

In general, it appears that previous experience--but more specifically experience with pre-school children and/or disadvantaged children--increases a teacher's ability to predict a child's achievement. Previous teaching experience also raises the correlation of teacher with observer in rating the same child. A curious exception is noted in that the sample of teachers with Masters Degrees did not follow the general pattern. It is also noted that the total number of years teaching experience (in which grade taught is not considered) is less consistent a determinant than the two variables of experience with disadvantaged children and experience with pre-school children. This suggests that there may be specific attitudes and judgements related to working with Head Start which are not related to general teaching experience.

It is unfortunate that the analysis of predictive success was limited to Center A. The results in Chapter V indicate differences between Center A and B. Center B had only four head teachers, two of whom had no previous teaching experience.

An appraisal of the naturalistic observations and Behavior Inventory data (see Chapters III and V) indicates that the structure and organization of the classroom (which ultimately depends on the teacher's philosophy and behavior) does affect the frequency of certain types of child behavior. For instance, the amount of aggression displayed by a child is partially a function of how much control and planning exist in the classroom, as well as the personality of the child. The paucity of material collected about the teachers and the small number of teachers made it impossible to pursue these questions further at this time. However, it is felt that more attention must be turned to differentiating the teacher variables which affect teaching performance and the ability to evaluate children's behavior consistently.

C. Future Plans

At present our particular area of interest lies in investigating the teacher's verbal behavior which acts as a stimulator to cognitive development. An observational-coding system is being designed which differentiates the cognitive from non-cognitive verbal behavior of teachers. The cognitive behavior is evaluated on two dimensions: one, the level of complexity of the basic cognitive structures used by the teacher in communicating with the children (such as indicating similarities and differences, focussing attention, problem solving) and two, the complexity of the cognitive response that the teacher demands of the child. Attention is also being given to evaluating non-verbal areas relevant to cognitive development, such as: type of control strategies and sanctions used in maintaining discipline, the organization and structure of the classroom program, and the use of non-verbal cognitive-stimulating equipment such as Montessori Materials.

A pilot study indicates that such an observational-coding system does differentiate between teachers and also shows differences related to type of activity. A methodological study is being planned to work out technical questions and to evaluate the reliability of observer-coders.

TABLE IV - 1

A COMPARISON OF PRODUCT MOMENT CORRELATIONS OF HIGH AND LOW TEACHER GROUPS WITH PREDICTIVE AND CRITERIA VARIABLES

I. Question of Teacher-Observer Agreement - Evanston & Woodlawn (N=44-80)

	Level of Education		# Years with Disadvantaged		# Years with Pre-school		# Years Teaching Experience	
	Master	College	High	Low	Some	None	High	Low
1. Summer Teacher Eval. of Kdg. Readi. vs. Observer-Achieve.	.58	.37	.60	.28	.61	.17	.53	.36
2. Summer Teacher Kdg. Adaptation vs. Observer-Kdg. Adapt.	.61	.36	.50	.42	.54	.31	.49	.42
3. Summer Teacher Achievement vs. Observer-Achievement	.44	.35	.41	.45	.47	.30	.38	.53
II. Question of Accuracy of Prediction (Evanston only (N=34-48)								
1. Summer Teacher kdg. Readiness vs. Fall Teacher Kdg. Readiness	.42	.56	.59	.32	.59	.28	.46	.52
2. Summer Teacher Adaptation vs. Fall Teacher Adaptation	.67	.47	.60	.52	.61	.47	.55	.63
3. Summer Teacher Achievement vs. Fall Teacher Achievement	.49	.54	.61	.42	.55	.46	.57	.35
4. Summer Teacher Achieve. vs. Reading Readi. on Report Card	.48	.71	.77	.39	.71	.46	.65	.54
5. Summer Teacher Achieve. vs. School Task Summary Report Card	-.49	-.48	-.52	-.51	-.48	-.55	-.51	-.51
6. Summer Teacher Behav. Inv.-Verbal Assertive. vs. Rept. Cd.-Verb. Ass.	-.29	-.58	-.59	-.26	-.58	-.18	-.43	-.41

CHAPTER V

NATURALISTIC OBSERVATIONS AS A TECHNIQUE FOR ASSESSING THE BEHAVIOR OF HEADSTART CHILDREN

Abstract: During the course of the Head Start program, 232 children were observed for twenty minute periods on the average of once every other day in Centers A and B. Observers were instructed to write complete narrative accounts of the child's behavior. A limited sample of these narratives on each child in the sample were coded in four separate sets of categories. These sets of categories were: 1) Behavioral Function (motive), 2) Level of Behavior, 3) Level of Social Participation, 4) Type and Level of Speech. The large majority of functional behavior was coded as Autonomous Achievement or Compliance. This is seen in terms of the failure of the coding system, for methodological and theoretical reasons, to differentiate actions within a stream of ongoing behavior. It is argued that future naturalistic observations of the motives and/or consequences of the child's behavior should utilize pre-coded techniques. Our findings suggest that two important dimensions along which to characterize the preschooler's behavior are first, the degree to which the child makes an effort to adapt the manner of his performance vs. the degree to which he is content to do things in familiar non-challenging ways; and, second, the level of his social interaction as measured by the Parten Newhall Index for Social Participation. The paucity of useful data obtained from categories for conceptualizing type and level of speech may be accounted for by an overburdening of our observers who were asked to attend to too much behavior at a given time, and could not report speech accurately enough for precise coding. Pre-coded ratings made by observers at the end of each observation were found to have good range and variability as well as being associated with other rating and test behavior. In sum, our methodological results suggest the need for situations in which an observer is asked to attend to a limited number of variables at a limited time and where he codes the observations directly into the category system.

A. Introduction

In addition to the use of test instruments and rating scales to assess the cognitive, social and emotional abilities of the preschool children in the study, naturalistic observations were included to obtain an objective account of the naturally occurring behavior of the children in the school settings of Center A and B. Our aim was to develop a set of behavioral categories which would enable an observer to evaluate the on-going behavior of pre-schoolers in categories which are useful predictors of his school readiness and social functioning.

This chapter was prepared by Richard Kramer

Further, since the intellectual and emotional functioning are so intimately associated in the young child, the evaluation of cognitive status as well as the assessment of social behavior in the school setting might be facilitated by such a set of categories.

The theoretical, methodological and practical issues involved in the use of naturalistic observations are myriad, aspecially with large populations. Initially, our attempt was to obtain objective non-biased accounts of the behavior of our classroom subjects. In view of the fact that there exist many preconceptions as to the difference between "disadvantaged" and "non-disadvantaged" groups, this project attempted to avoid constructing our results by asking observers to record behavior rather than to precode it. Still, while they were asked to be as descriptive as possible, it is clear that it is theoretically impossible to avoid categorizing behavior through the use of language even without the use of set categories for pre-coding. Our aim was to apply a system of categories to the observed behavior and to determine which, if any, categories were applicable to behavior in a nursery school setting. From a set of categories which were applicable, i.e., into which behavior could be coded, a further set of categories was sought which could be used to differentiate amongst children and which might predict test behavior and/or teacher ratings.

It is best to use a small set of categories in so far as pre-coded observations are most useful with large populations, and methodologically the aim of our project was to develop such a shortened list. However, inference based on our substantial findings must be tendered with qualification for two reasons. First, the low level of our observer reliability and coder reliability attenuate certain of the substantive findings. Secondly, although our original list of categories was held to be a set of mutually exclusive and exhaustive set of categories, we have no information with respect to how a behavior classed as X might have been coded if category X were not a coding alternative. Our list was taken to be exhaustive in order to avoid making decisions as to which behavior to leave uncoded (exception of uninvolved behavior). Therefore, many behaviors coded in non-applicable categories or in categories whose scores had essentially no variance might, under other circumstances, contribute to the scores which were used in our substantive analyses. To the degree that this happened, our substantive findings are inconclusive.

It is emphasized that this N.O. research must be seen as an intermediate step. Our aim was to develop a set of categories which might be used in observing children and which would be sufficiently refined to produce variance in scores to differentiate between children and to relate observed behavior to teachers ratings, test performance and school achievement.

B. Design

At the beginning of the Head Start program, a group of ten observers (graduate students and housewives with college degrees) were stationed in Center A (one in each of ten classrooms). Having been given a two-day seminar in the technique of naturalistic observation, they spent the first week of the program [illegible] refining their skill. Commencing in the second week, each observer made observations of 20 minutes on each of eight of the fifteen children in his classroom each day. This schedule was carried out for periods ranging from five weeks (in five classrooms)to six weeks (in five classrooms). Since there were fifteen children in a classroom, each child was observed on the average of once every other day. The school day was divided into eight 20 minute periods for the observations, and the time of day during which a child was observed was varied to insure an equal number of observations on him during each time period through the course of the program.

During each 20 minute observation period, the observer recorded 15 time samples of behavior. Each time sample involved 15 seconds of looking at the child and 45 seconds of recording his behavior. The five extra minutes were utilized for finding the child, making preliminary description of the setting, and making four behavior ratings at the end of the observation.

In Center A, 150 children were observed. Half of them were observed for 11-13 20 minute periods each, and half were observed for 14-16 20 minute periods. Five observers went to Center B during the final two weeks of its program where some 75 children were observed for 5-8 20 minute periods each.

The observers were instructed as follows: The purpose of the observation procedure is to obtain an objective account of the child's behavior in a naturalistic setting. Our guiding concern is to find out as much as we can about the everyday activities that are natural to the child. In investigating behavior of the child, we are interested in what he does on his own and what he does in response to the nursery school setting.

Our guide line for the observations is that they must be objective, complete and legible...The observations themselves should be akin to the motion picture form of a child's behavior--unedited, panoramic, and in fine, very sharp focus...

The Behavior Rating Form (Appendix 1) will be used. When you are ready to observe a given child, fill out the top of the form with the date, the child's code number, the time begun, and the number one in the blank next to the word 'page'. Then describe the situation and setting in the appropriate box. This box will not be used in subsequent Time Samples on the child unless the general situation (including major group activity, major child activity, physical setting, etc.) changes. Having filled this out, begin your observation of the child. The STIMULUS box will only be filled in when relevant (e.g. gets hit by another child). The BEHAVIOR box should include the specific acts the child is engaged in, his interaction with others, a description of the objects he is dealing with, and an account of his manner of manipulation of these objects. The EXPRESSION box gives a place to indicate in adverbial phrases, the 'how' of the child's behavior as well as a description of his facial and/or bodily expression.[1]

---

[1] These are excerpts from the observers' guide. Observers were also given a section from Barker, Kounin and Wright's The Midwest and its children. as a further guide to observation.

C. The Coding System

Two primary codes were given to each of the 15 time samples in a given protocol. They involved a categorization of the function of the child's behavior, i.e., his "intent" or "motive", and a scoring of the "level" at which the child's intent found behavioral expression. In addition, the child's play was coded according to the Parten and Newhall Social Participation Index (1943), his speech was coded in terms of its egocentricity vs. its social adaptive character, and counts were made of the number of times the child was laughing or smiling.

1. Function

In developing coding categories for the immediate function or behavioral motive of the child's actions, we relied almost totally on already existing category names, occasionally refining the definitions, extending or delimiting the scope of the categories involved. The two exceptions were the categories of 'resistance' and 'compliance'. In each of these cases, the category was split in two. Compliant behavior geared to developing one's own skills through behaviorally adopting the suggestions of another was differentiated from compliance engaged in solely for the sake of forestalling negative sanctions or avoiding punishment. Likewise, resistance geared to allowing the self greater autonomy and self-directedness for the goal of learning was distinguished from resistance engaged in for the sake of maintaining one's status or power with respect to the person requesting modification of one's behavior. Simply stated, we had hoped to distinguish between the child who says, "No, I want to do it my way." and the child who says, "No, I refuse to do anything that you ask of me."

The set of sixteen categories which were designed to be mutually exclusive was as follows:

a. Aggression upon another
b. Aggression upon self
c. Autonomous Achievement
d. Avoidance-withdrawal
e. Compliance A
f. Compliance B
g. Control Dominance
h. Friendship-affiliation
i. Giving help to another
j. Giving nurturant attention to another
k. Resistance A
l. Resistance B
m. Securing help for self
n. Securing nurturant attention for self
o. Uninvolved (no level score)
p. Uninvolved, watching others (no level score)

(Note: The exception to the mutually exclusive rule for coding was in the case of aggression upon another. Any behavior expressed in aggressive form where the primary motive was not aggression was given two codings.)

2. Level

With the exception of functions 15 and 16, each time a unit was scored for behavioral function, a judgement was made as to the developmental level at which the function found expression. It was felt that the same motive could be expressed in more or less complex ways at higher and lower levels of social interaction. This scoring was to facilitate greater discrimination between children who express their autonomy, aggression, compliance, avoidance, etc. in significantly different ways. For example, level one stressed the fulfillment of bodily needs while level six (virtually never used in this pre-school population) stressed the modifcation of the self's social role.

The following is a list of generic descriptions of each of the levels:

Level 1. The goal of the behavior is the fulfillment of bodily needs of self or another and/or the goal is served through the use of the body and bodily contact, or the failure to do so.

Level 2. The goal of the behavior is the fulfillment of object needs of self or another and/or the goal is served through the use of objects, or the failure to do so.

Level 3. The goal of the behavior is the fulfillment of activity needs of self or another and/or the goal is served through the performance of some activity, or the failure to do so.

Level 4. The goal of the behavior is the fulfillment of needs for adaptation in the manner of performance of some activity and/or the goal is served through an adaptation in the manner of performance of an activity, or the failure to do so.

Level 5. The goal of the behavior is the fulfillment of needs to perform adaptively with respect to other people and/or the goal is served through an adaptation in the manner of the self's behavior toward other people, or the failure to do so.

Level 6. The goal of the behavior is the fulfillment of needs to modify or extend the self's social role and/or the goal is served through a modification or an extension of the self's social role, or the failure to do so.

Descriptions of behavior at each of the levels were made separately for each of the motive categories, with the exception of the two uninvolved categories. As examples of categories and of the descriptions of the behaviors at different levels, the sections from the coding system detailing the rules for coding Autonomous Achievement and Compliance A are given in full.

Autonomous Achievement: The goal of the behavior is to be independent of help and/or opinions of others, either through initiated activity or in performing some activity that is requested by another. Help is sought only for enabling the self to continue independently of further help. The primary goal is autonomous achievement. Usually, though not always, such behavior is characterized by the child's taking the initiative. If the achievement is in response to a request, code here only if the child utilizes the stimulating conditions to perform independently for himself. In a free play situation much behavior is aut. ach. It is not necessary that the child actually complete a task to code aut. ach.

Level and Content

1. Initiates attempt to meet bodily needs or perform routine tasks. The situation is such that the behavior is not simply Compliance (A) with what is being asked implicitly by the procedure of the classroom. May ask for permission but not help.

2. Initiates attempt to get materials for activity by self. May ask for permission but not help.

3. Persists at and/or completes a task. The goal is finishing activities non-body need related that are assigned. Instrumental behavior includes attention to task or activity and resistance to being distracted. The child continues working at an assigned task until it is completed, does not ask for help with surmountable obstacles.

4. Selects and initiates own activity. Does not wait for suggestions, urgings or invitations or leadership from others. Initiates general conversatio (not friendship-affiliation). May ask for permission, but not help; may ask for instruction so that he may continue unaided.

5. Facilitates ongoing activity or activity in the process of being initiated by group or self in group. Plays the role of mediator, compromisor, trouble shooter, provider of information or assistance, acquiring materials necessary to the activity. The goal is achievement (getting the job done) and NOT giving help, or aggression or giving or getting nurturance.

6. Seeking to remove frustrating non-physical (social) barriers to goal of self-initiated activity. Does not ask for any assistance. The barrier is large enough for the child to reasonably be asking for help.

Function

Compliance A: The goal of this behavior is deriving satisfaction through the enactment of the suggestions of another. The individual may comply in order to learn, he may comply in order to gain a feeling of acceptance of self (perhaps through being accepted by another). Asking for directions and complying with suggestions about how to do something may function to enable learning. Complying with the teacher's request to wash one's hands may function to enable the individual to feel accepting of self in that he may interpret the teacher's acceptance of his compliance as meaning he is a good boy, a worthy boy.

Level and Content

1. Complying with verbal or gestural requests of another anent bodily matters.

2. Complies with verbal and/or gestural requests of another to get an object for and/or give an object to another.

3. Complying with verbal or gestural request of another to perform a certain activity or to not perform a certain activity.

4. Complying with the verbal or gestural request of another as to how to perform a certain activity.

5. Complying with the verbal or gestural request of another as to how to behave toward others (or how not to)

6. Complying with verbal or geatural request of another as to how to facilitate an ongoing group activity. This level is hard to code given but one action. If the teacher asks that a child be her helper, and he complies, it would be coded here. Since individual actions themselves might not indicate that the individual has adopted some role or behavioral pattern at the request of another, code here only when there is clear indication that this is the case.

3. Social Participation

This code represents a summary rating of the behavior throughout a time sample. The highest appropriate number is assigned for the whole sample.. Interpret the concept of "play" in a broad way to include all activity, which may be cognitive, free-play, conversation, etc.

a. Unoccupied Behavior: The child apparently is not playing at all, at least not in the usual sense, but occupies himself with watching anything which happens to be of momentary interest. When there is nothing exciting taking place, he plays with his own body, gets on and off chairs, just stands around, follows the teacher, or sits in one spot glancing around the room.

b. Solitary Play: The child plays alone and independently with toys that are different from those used by the children within speaking distance and makes no effort to get close to or speak to the other children. His interest is centered upon his own activity and he pursues it without reference to what others are doing.

c. Onlooker Behavior: The child spends most of his time watching the others play. He often talks to the playing children, asks questions, or gives suggestions, but does not enter into the play himself. He stands or sits within speaking distance of the group so that he can see and hear all that is taking place. Thus he differs from the unoccupied child, who notices anything that happens to be exciting and is not especially interested in groups of children. Only code if there is evidence of some attempt to be related to the group - by speech or move close.

d. Parallel Play and Activity: The child plays independently, but the activity he chooses naturally brings him among other children. He plays with toys which are like those which the children around him are using, but he plays with the toys as he sees fit and does not try to influence the activity of the children near him. Thus he plays beside rather than with the other children (cf. solitary play).

e. Associative Play and Activity: The child plays with other children. There is borrowing and lending of play material; following one another with trains and wagons; mild attempts to control which children may or may not play in the groups. All engage insimilar if not identical activity; there is no division of labor and no organization of activity. Each child acts as he wishes, does not subordinate his interests to the group. A conversation between children or child and adult is considered Associative Activity.

f. Cooperative or Organized Supplementary Play and Activity: The child plays in a group that is organized for the purpose of making some material product, of striving to attain some competitive goal, of dramatizing situations of adult or group life, or playing formal games. There is a marked sense of belonging or not belonging to the group. The control of the group situation is in the hands of one or two members, who direct the activity of the others. The goal as well as the method of attaining it necessitates a division of labor, the taking of various roles by the different group members, and the organization of activity so that the efforts of one child are supplemented by those of another.

4. Speech Categories

a. Pure egocentric speech (repetition, action monologue, vocalized self-guidance, etc.) and egocentric speech in social situations (collective monologue, group guiding monologue, etc.).

b. Socialized speech in egocentric style (attention getting questions, neologistic commands, etc.).

c. Socailized speech (answers to questions, didactic statements, etc.)

5. Expression

a. Smiling

b. Laughing

In addition to these four codes that were made on each time sample, the observer was asked to rate the child, based on his twenty minute observation, on a 1-7 scale on each of the following four items.

a. Appearance: How does the child look? (7 point scale from "very near, clean and polished", to "very unkempt and dirty".)

b. Motor coordination: How does the child handle his body? (7-point scale from "excellent balance and coordination", to "very clumsy and awkward".)

c. Speech: How often did the child verbalize during the observation? (7-point scale from "very often", to "rarely".)

d. Engagement: Please rate on a 7-1 scale your reaction to the child's overall engagement in the nursery school setting during the observation period. You should take into account whether or not the child seems generally able to utilize the opportunities for play, interaction and learning that surround him. It somehow seems natural that a child rated at 7 would be happy in what he was doing while a child at 1 would be fairly withdrawn.

D. Results

1. Reliability

a) Coder Reliability

A set of observations for two children (15 time samples each) was coded by each of seven coders. The coding for each time sample was compared with a criterion code which was established on the basis of consensus. The coding supervisor was used as a standard when necessary. The number of disagreements from the criterion code was calculated. Table V-1 presents the number of disagreements in each coding category for the two protocols. There were seven coders coding 30 time samples in 6 ways. For each of the six ratings there were 210 judgments. Summing over the six ratings there were a total of 1260 judgments. Total percent agreement across all coding categories was 75%.

b) Observer Reliability.

In order to check on the reliability of the observers, it was arranged to have children simultaneously observed by two observers. These observations were then coded blind by a regular coder whose intracoder reliability was high.

Several problems presented themselves in checking for observer reliability. First, there were mechanical difficulties, such as the problem of synchronizing the beginning of each time sample, which reduced the number of agreements in scoring. Also, the observers had difficulty hearing and seeing the child identically because of different placement in the room. A second difficulty was the variation in the amount of elaborative detail which different observers included in the protocols. The elaboration often affected the coding, especially if it involved information about teacher instruction or interaction with other children. The observers were not aware of the coding system and the terse observers often did not give sufficient information to make distinctions called for by the coding system.

Third, the application of the coding system rules maximized some of the inherent discrepancies. For instance, the first item of behavior in the time sample was chosen as the unit to be coded. This rule was applied in order to

TABLE V - 1

DISAGREEMENTS FROM CRITERION CODES AND PERCENT AGREEMENT (FOR SEVEN CODERS CODING 30 TIME SAMPLES)

	# Disagreements	% Agreement
Function	76	64
Level	62	70
Participation	41	80
Speech I	38	82
Speech II	71	66
Expression	23	89

$$\text{Total Reliability} = \frac{\text{\# agreements}}{\text{Total \# of judgments}} = 75\%$$

reduce variation caused by differences in coder qualitative judgments as to which behavior in a time sample was primary. However, this rule of first item maximized differences due to the observers' lack of synchronization. For these reasons, the observer reliability information using function and level scores has been analyzed in two ways. All units per time sample instead of just the first were coded. All Function-level combinations in each of the two matched time samples were identical for 51% of the time samples. However, in 83% of the time samples there was at least one agreement in the combinations.

These figures indicate that sufficient variability occurs to warrant changes in observer training in future. The greatest difficulty is delimiting the relevant behavior unit to record. In general, observers saw the same behavior but noted more or less of the peripheral environment, and were more or less able to reconstruct the behavioral sequence. This problem might best be controlled by focusing the observers' attention on the type of information which is needed to make the coding discriminations.

2. Behavioral Function Categories

Twenty minute observations totaling 15 time samples of behavior were taken on each child at the rate of once every other day for periods ranging from four weeks in Center B to seven weeks in Center A. Of the total set of observations on each child, only three to four observations, spaced at equal intervals throughout the period of weeks during which the children were observed, were coded for each child. The range of the number of time samples coded for each child was from 45 to 60. The percentage of time samples coded scored in each functional category were computed for each child using as a base the number of time samples in which he had been observed. Each child had a Percentage Fumction Score for each function and the sum of all these Function Scores for each child was 100. Table V-2 presents the number of children in the sample of 232 whose Percentage Function Scores fell in the ranges defined in the table for each function.

The percentage scores in only two categories (Autonomous Achievement and Compliance A) reveal much variance among children. It appears that the greater majority of all behavioral units were scored in either one or the other of two categories. (Aut. Ach. and Comp. A) This result is the function of many factors, but primarily it suggests that the network of categories employed to codify the naturalistic behavior of our subjects was too poorly refined to discover much more than whether the child was doing something on his own or in response to the suggestions and demands of the school setting. In a sense, Aut. Ach. and Comp. A proved to be 'catch all' categories. It cannot be concluded from this result that our subjects were unaggressive, non-controlling, unfriendly, and non-nurturance seeking. That these categories only received minimal use reflects inadequate observer preparation and an unwieldly coding design, especially in view of limited time and number of observations to be coded. This points to the necessity for pre-coded observations with fewer and more clearly operationally defined coding categories to be used by the observer on a given day.

TABLE V - 2

DISTRIBUTION OF PERCENTAGE FUNCTION SCORES FOR EACH BEHAVIOR FUNCTION

%	Number of Subjects Given Each Percentage Score For the 16 Behavior Functions							
%	0-9	10-19	20-29	30-39	40-49	50-59	60-69	70-79
Agg-An	226	6	0	0	0	0	0	0
Agg-Self	232	0	0	0	0	0	0	0
Aut Ach	3	22	54	60	61	21	7	4
Av. Wd.	225	5	1	0	1	0	0	0
Comp A	4	37	52	54	52	22	9	2
Comp B	213	15	4	0	0	0	0	0
Con.Dom.	226	6	0	0	0	0	0	0
Friend	212	18	2	0	0	0	0	0
Give H.	231	1	0	0	0	0	0	0
Give A.	232	0	0	0	0	0	0	0
Res A	215	14	3	0	0	0	0	0
Res B	224	8	0	0	0	0	0	0
Sec H	228	4	0	0	0	0	0	0
Sec A	184	42	6	0	0	0	0	0
Unin	214	15	2	1	0	0	0	0
Unin W	194	34	4	0	0	0	0	0

N = 232 including all Middle and Working Class Children in Centers A and B

A Percentage Function Score of 0 means in all of the time samples coded the child never expressed this function. A score of 21 means that in 21% of the time sample coded, he expressed this function.

While the minimally used categories may prove to be the most relevant, the fact that we coded only a maximum of 60 time samples out of an average total of 160 (four observations taken at equally spaced intervals throughout the program were coded for each child) has seriously depressed the range and variance of these scores. Nonetheless, a sufficient number of children had more than 10 percent of their behaviors scored in almost all of the minimally used categories to warrant analysis of the association of these scores with other test and rating data.

One major source of error in this type of naturalistic observation and coding is due to the adult-centered nature of the coding categories. If the child were to articulate his motives he would probable not make the same distinctions regarding his behavior as an adult observer. That is to say, as the child moves through his environment, he does not compartmentalize his intentions. He is not at one moment intending to be aggressive and at another moment intending to form a friendship or intending to withdraw from an activity. For the child, the ongoing behavior, his day's activity, is geared toward a set of personal ends which he cannot articulate. The adult observer artifically distinguishes one act from another, and in so doing also artificially distinguishes one motive from another. There is a viable rationale for so doing in terms of noting the consequences of the child's behavior, but motives are imputed. Regardless of how well a set of observations are taken and coded, the dimensions along which we characterize the child's behavior are products of our own preconceptions about that behavior. For this reason alone, there is bound to be a great deal missed in such a naturalistic observation format.

Further in reflecting upon the naturalistic setting in which the observations were taken, it has become clear that in large measure the categorization of a given piece of behavior is more a function of the school setting than of the child's intentions. One example illustrates this important point. In Center A, the playgrounds were filled with equipment; slides, jungle-jims, tricycles and smaller toys were in abundance. In Center B, the playground was an empty lot. As the child enters the setting, he is ready to behave, and his behavior will be structured at least in part, by his environment. Let us call the latent behavior that a given child is about to display x. X can be thought of in terms of an unconscious motive or a physiological propensity toward motor expression; it makes no difference The point is this. In Center A, x might be expressed in climbing a jungle-jim and shouting "I am on top!" In Center B, the same x may be expressed (for lack of jungle-jims) by the child pushing another down, steppin on him and shouting "I am on top!" In Center A, the behavior is coded as Autonomous Achievement, in Center B as Aggression.

Two conclusions may be drawn from this example. First, comparison of children in different settings becomes impossible. Second, categorizing behavior in terms of social motive or consequence in some cases says less about the child than it does about the setting.

3. Level Scores

For each child, the percentage of total coded time samples scored at each of the six levels was computed (similar to the computation of Percentage Function Scores). Table V-3 presents the number of children in the sample of 232 whose Percentage Level Score fell in the ranges defined by the table for each level.

TABLE V - 3

NUMBER OF SUBJECTS GIVEN EACH PERCENTAGE LEVEL SCORE

%	0-9	10-19	20-29	30-39	40-49	50-59	60-69	70-79
I	21	62	71	51	16	6	5	0
II	202	26	4	0	0	0	0	0
III	3	50	66	49	48	14	2	0
IV	3	26	52	58	50	28	14	1
V	228	4	0	0	0	0	0	0
VI	227	5	0	0	0	0	0	0

a N = 232 including all Middle and Working class children in Centers A and B

These results indicate that within the limited age span of pre-school children it is possilbe to apply the notion of developmental level to a categorization of the child's behavior. It seems clear that levels 5 and 6 are inapplicable for the most part, perhaps partially because of the level of inference involved in rendering a code of 5 or 6, but probably because of the general level of maturity of the pre-schooler and of the kinds of social demands made upon him in the school setting.

Level 2 seems conceptually poor by virtue of the fact that "object needs" are not clearly distinguishable from "activity needs" which involve objects. Levels 1, 3, and 4, show sifficient range and variance to warrant close scrutiny. It appears that the subjects' behavior, no matter what the function, was characterized in one of three ways. Either he was concerned with the performance of bodily functions (many times as a result of direct commands by the teacher), or he was characterized as fulfilling the need for activity, or he was characterized as fulfilling the need for adaptation in the manner of his activity.

These three levels seem to be distinct, and seem to be important in terms of characterizing the level of behavior in pre-schoolers. How much does the child's behavior center around bodily concerns, how much does he move about and seek activity, and how much does he attempt to adapt his performance or learn new ways of doing things. While our results are heavily attenuated by relatively poor observer and coder reliability, it seems apparent that the distinction between level 3 behavior and level 4 behavior is a viable and important dimension along which to observe the behavior of the pre-school child.

4. Social Participation

Table V-4 presents the number of children in the sample of 232 scoring different percentages of total responses in each of the six Parten and Newhall categories of Social Participation.

TABLE V - 4

NUMBER OF CHILDREN GIVEN EACH PERCENTAGE SCORE FOR THE 6 TYPES OF PLAY

%	0-9	10-19	20-29	30-39	40-49	50-59	60-69	70-79	80-89
Unocc.	164	48	16	3	0	1	0	0	0
Solitary	91	77	38	18	5	3	0	0	0
Onlooker	231	1	0	0	0	0	0	0	0
Parallel	6	29	45	55	41	35	15	4	2
Assoc.	1	22	43	58	44	35	18	8	3
Cooper.	232	0	0	0	0	0	0	0	0

N = 232 including all Middle and Working class children in Centers A and B

It appears that Cooperative play is an infrequent characteristic of this age group, and that Onlooker behavior is to be a special category of Unoccupied Behavior. The other categories seem valuable in characterizing the level of social participation, and these play categories can be used to distinguish between episodes of play in the preschool classroom.

5. Speech Categories

Table V-5 represents the mean number of time samples out of a total of from 45 to 60 that were coded for each type of speech.

TABLE V - 5

SPEECH

Type of Speech	Mean Number Time Samples Coded
Egocentric Speech	7.61
Socialized Speech in Egocentric Style	8.30
Pure Socialized Speech	16.02

As the coding rules called for the scoring of direct quotes only, much speech was left uncoded. The paucity of speech codes seems more a result of the tremendous difficulty arising when observers are asked to maintain detailed records of speech while at the same time attempting to write a complete narrative account of the child's activity in the time sample, than it does to any failure of the coding categories to distinguish types of speech. Precoded observations would seem to alleviate this kind of pressure on observers, and there is no clear indication that egocentric vs. socialized speech is not an important dimension along which to assess the functioning of the preschool child.

6. Expression

The mean number of time samples coded for smiling or laughing was 14.72. Again, this is out of a possible 45 to 60 time samples.

7. Observer Ratings

At the end of each 20 minute observation period, the observers made four ratings of the child's total performance. The ratings for each child were averaged across the 5 to 16 available ratings to provide a single mean score for each child. Table V-6 represents the distribution of mean scores for the four observer ratings.

TABLE V - 6

NUMBER OF CHILDREN GIVEN EACH MEAN SCORE FOR EACH OF THE FOUR OBSERVER RATINGS

Rating	N	.1-1.0	1.1-2.0	2.1-3.0	3.1-4.0	4.1-5.0	5.1-6.0	6.1-7.0
Appear.	230	1	0	10	68	104	39	7
Motor	232	0	0	2	57	99	63	8
Speech	232	0	42	61	62	481	15	1
Engage	232	0	0	8	51	105	63	2

The distributions on all four of these ratings suggest that the precoded method will prove most effectual in terms of gaining variability to discriminate among children. That the distribution of scores for the Engagement rating seems to approximate a normal one is expecially interesting in terms of the item the observers were asked to rate upon. Essentially, they were asked only for their gut reaction to how well the child was engaged in and utilizing and profiting from the school setting.

8. Relationships to Other Data

A number of the scores derived from the coding of the naturalistic observations and selected for their greater range and variance were correlated with subsequent school performance, achievement and reading and number readiness tests administered by the school system in Center A during the Spring following the Head Start session. Binet scores taken during the Head Start summer program were also included.

The scores derived from the naturalistic observations which were correlated with I.Q. and school performance were as follows:

a) The percentage scores for all the functional categories. Some categories were combined where the range of separate scores was limited and where it was felt the categories included similar behavior, such as Securing Help with Securing Attention.

b) Mean level scores for all behavior scored in a given function. Each child received a Mean Level Score for each function category in which three or more time samples had been coded.

c) The means of the observer ratings on Appearance, Speech, Motor Coordination and Engagement in the school setting.

The scores used for correlation with scores from the naturalistic observations were:

a) The national percentile on the Reading and Number Readiness

b) A factors score indicating school performance based upon school report cards (see Chapter VII)

c) I.Q. as measured by the Stanford-Binet during the Head Start session.

Table V-7 presents all significant correlations between these scores and some of the scores derived from the naturalistic observations and the observer ratings. More information about the usefulness of these scores in predicting school achievement may be gained from Chapter VII.

TABLE V - 7

SIGNIFICANT (.02) CORRELATIONS OF NATURALISTIC OBSERVATION SCORES WITH I.Q., ACHIEVEMENT TEST, READING AND NUMBER READINESS TEST, AND REPORT CARD SCORES

	I.Q.	Natl. Percentile Rank Ach. Test	Report Card School Performance
Mean Level Auton Ach.		.252 (N=72)	
Mean Level Securing Help + Securing Atten.	.293 (N=105)	.312 (N=72)	.321 (N=72)
Mean Observer Rating for Motor Coordination	.225 (N=180)	.291 (N=72)	
Mean Observer Rating for Amount of Speech	.344 (N=180)		.315 (N=68)
Mean Observer Rating for Engagement in School Setting		.303 (N=72)	.331 (N=72)

Note: This is a sample of both middle and working class children from Center A only

None of the Percentage Function Scores correlated significantly with either the I.Q. scores or the subsequent achievement, readiness, and performance scores. However, the Mean Level Score for Autonomous Achievement (the most frequently used single category) and the Mean Level Score for the combined categories of Securing Help and Securing Attention did show some relationship to the test scores.

Methodologically, these findings suggest that global scores based on ratings of gross segments of behavior predict cognitive ability and performance as well if not better than more refined scores of behavioral motives. Further, the scores for the level at which a child plays or seeks help and attention bear more relationship to his cognitive ability and academic performance than do the scores for the functions of his behavior. These findings suggest the viability of using global and behavioral level ratings as 'predictors' of cognitive ability and school performance.

Substantively, these findings would seem to indicate that the over-all assertiveness of the child in terms of his use of speech, his lively participation in activity and his active search for new activities and for help and attention in dealing with them is the most important dimension along which to evaluate him if one wishes to predict school readiness and cognitive ability.

E. Summary and Recommendations

Our experience in attempting to use an extended and complex format for naturalistic observation on a large population of pre-school children has brought to light many theoretical and methodological problems; the format used has posed innumerable practical problems in terms of time and expense.

Our findings suggest that when observers do not have knowledge of the coding system, and when that coding system is utilized in an attempt to codify behavior from a number of different perspectives (function, level, social participation, type of speech, expression) and when it is impossible for reasons of time and expense to code more than 60 minutes fo behavior over a seven week span, it is extremely difficult to gain data that is good enough to assess either the viability of the categories employed or the functioning of the children themselves.

Succinctly, our findings with respect to behavioral functions suggest that precoded observations are necessary for two reasons. Firstly, they eliminate the time consuming task of coding raw observations, and the slippage involved, and secondly, they would allow for a more complete viewing of the child throughout a given time period, in this case seven weeks.

Our findings suggest that an important dimension along which to view children is the extent to which they spend time attempting to adapt their own manner of performance as opposed to the time spent playing in familiar ways (i.e. levels 3 & 4). Further, the Parten and Newhall categories of Social Participation seem relevant and applicable to this population of children. These categories could easily be adapted to precoded observation methods. Further, the more global ratings of the child's speech and engagement in the school setting were the best "predictors" of I.Q. and school performance.

In summary, naturalistic observation is an expensive and time consuming task. Our recommendation is that with large populations a number of methodological and practical problems can be circumvented by the use of precoded observations. Further, a format which calls upon the observer to focus upon a limited number of variables at a given time would seem to allow for better quality data with which to assess the functioning of the pre-school child.

CHAPTER VI

ASSESSMENT OF THE COGNITIVE AND SOCIO-EMOTIONAL HOME ENVIRONMENTS OF LOWER-CLASS HEAD START CHILDREN.

Abstract: This aspect of our research project was directed toward at least two issues: How useful are maternal variables in predicting Head Start children's academic achievement? and What is the influence of different kinds of measures of these variables on this prediction. As a result, in this program a variety of types of instruments were administered to the mothers, including a 2 ½ hour Parent Interview, constructed for this purpose; a Checklist Behavior Inventory, designed by the Office of Economic Opportunity for use with teachers, and adapted for the mothers; an Education Attitude Survey adapted from another preschool project of Hess and Shipman; Witkin's Embedded Figures Test; Draw-A-Person; and James Bieri's Cognitive Roles Test. From most of these instruments a variety of maternal variables were assessed.

Generally, maternal variables do not predict Head Start children's school achievement as well as teachers' predictions and previous cognitive test performance scores. Three variables were found useful for further exploration: a) Educational Aspirations -- hoped for and expected, b) Openness of the Mother to Answer Her Child's "Difficult" Questions, and c) Frequency of Inperative Statements made by the mother--that is, the degree to which she gives commands to her child without the accompanying rationale for these directions. These variables correlated significantly with children's Stanford-Binet intelligence scores. With regard to changes in these children's behaviors, particularly Preschool Achievement Test scores, mothers' cognitive performances on certain perceptual and differentiation tasks were predictive of children's potential for change in level of performance.

With regard to instrumentation, certain measures are more practical, though less refined, than others. For a brief estimate of the maternal home environment we recommend use of a) the Education Attitude Survey or similar such 5-point checklist inventory, and b) Witkin's Embedded Figures Test or similar cognitive task, in addition to obtaining the basic demographic data. If, however, a more extensive investigation of the gome environments of these children is planned, an interview constructed around interests of the child, from the mother's cognitive frame of reference is essential. For example, the best information about the mother's personalization of her child is not obtained by asking her directly how her child differs from other children in the family, but possibly indirectly through a description of the various eating habits of the children.

Further analysis of these variables, and their relationships, is planned, including the development of rating scales for the total parent interview.

This chapter was prepared by Diana Slaughter

A. Rationale

The purpose of this section is to describe the approach of this research project to the study of the cognitive and socio-emotional home environments of some Chicago area Head Start children. For several reasons, we considered the design and evaluation of techniques for the study of the disadvantaged family crucial to our research program. First, the intercorrelation between familial structure, attitudes, and behaviors, and children's cognitive and personality development, has been consistently documented in the research literature. Extensive study in the area of parental behavior and child development, however, has been handicapped by the paucity of conceptual or theoretical appreaches to the study of the family as a socio-psychological entity. Only recently have such formulations begun to emerge. It is important, however, that as these formulations appear we consistently check and recheck our hypotheses regarding the role of the family in the growth of the children. Head Start populations provide a hitherto unavailable source of validation; too often in the past we have relied solely upon the children of university or college personnel enrolled in university nursery schools, for example, for testing hypotheses about significant precidtors of academic and social achievement in children.

Second, there are certain specific advantages to doing research with disadvantaged families. Children do not grow up in a vacuum. The attitudes and behavior of parents do determine in large part those of the children--we know this now. Since much of the lower-class child's lack of preparation for the school situation may stem from the quality of his home experiences, any intervention programs taken to facilitate the adaptation of this child to the present school system must consider what the home background experiences of these children are and have been. For an understanding of the longitudinal growth and development of these children, study of the familial background is essential: variables relatively unimportant for prediction of academic achievement level at earlier stages of these children's development, for example, may later assume considerable significance for actual performance.

Third, the problem of creating appropriate assessment techniques for family research is even more critical with lower-class populations. This, too, has been documented recently. If, however, we cannot communicate with their parents satisfactorily, then can we expect to communicate with these children? As we conceived our research, one of our goals was to investigate thoroughly some of the problems of assessment of relevant parental variables in this population. Ideally, this would result in a new instrument (or instruments) specifically found to be useful in the study of parent-child relations in lower-class populations. At the very least, specific recommendations for such instruments could be made.

B. Assessment Procedure

The initial problem of this research was the development of instruments and measures appropriate for obtaining information from lower class mothers regarding their children. In brief, what kind of information-seeking techniques would be most useful in predicting children's current and long-term functioning, particularly within the academic setting? As a result, we were not satisfied with the use of just one type of instrument, but instead experimented with several.

In this section we sill describe these instruments, and in the following section on data analysis we will report briefly some of our conclusions of this exploration.

The list of instruments utilized included:

1) An open-ended parent interview, administered during the summer of 1966.

2) A Behavior Inventory. This was primarily a checklist instrument adapted for use with the lower-class mothers to compare their observations about their children with those of teachers and observers. The Behavior Inventory, originally developed by OEO in Washington, includes such items as:

a) My child talks eagerly to adults about his own experiences and what he thinks.

b) My child appears to trust in his own abilities.

c) My child has a tendency to discontinue activities after exerting a minimum of effort.

Each mother was asked to rate the extent to which these and similar qualities are similar or dissimilar to her own child. Thirty-eight such items were included in this instrument. Twenty of these items referred very specifically to the child's social and cognitive behaviors within the school. Included also was an over-all checklist for evaluation of each child's readiness for kindergarten. Each mother was asked to rate her child's preparedness for this early school experience. In addition, mothers were asked to predict the level of their child's adaptation to the total school experience as well as his over-all achievement standing. Finally, mothers were asked to rate their child on his level of speech development, motor coordination, physical appearance, and probable overall engagement or involvement in kindergarten. All ratings were completed on a seven-point scale, with the exception of the five- point Readiness for Kindergarten scale.

3) An Education Attitude Survey. This instrument was composed of 26 attitudinal items. For each item the direction and intensity of the mothers' agreement or disagreement was assessed. The survey was originally developed as part of another ongoing research project at this center (Hess and Shipman, 1965) for assessment of the lower-class mother's attitudes toward the school system and found to be related highly to a number of measures of maternal and child behaviors. The scale ranged from Strongly Agree = 1 to Strongly Disagree = 5. Sample items include

a) The teachers expect the children always to obey them.

b) The only way that poor people can raise the way they live is to get a good education.

c) The best way to improve the schools is to integrate them.

d) I can do very little to improve the schools.

4) The Draw-A-Person Test. All mothers were asked to Draw-a-Person. We anticipate using this material for assessment of some of these mothers' cognitive and socio-emotional characteristica as a group. Such information, it was thought, could be a useful index to continuing latent problems in the homes of some children. It could also be another index of the cognitive complexity of individual mothers.

5) Witkin's Embedded Figures Test. This is a perceptual test we used which was developed by H. Witkin for use in the study of the relation of parental experiences to cognitive development in children. To date this instrument has been used primarily with middle-class respondents. Possibly, however, we hypothesized, the lower-class mother's ability to distinguish form from color and thus solve the perceptual problem presented by this task (i.e. finding a simpler hidden figure in a more complex design) will be related to her child's test and observable behaviors. If so, this would provide a clue to the kind of psychological instruments we could successfully employ for diagnostic purposes with this group. We followed the same procedure reported by Witkin in which only six of the original twenty-four cards were administered.

6) Bieri's Cognitive Roles Test. This was another instrument which we included in our psychological battery, primarily to investigate whether the lower-class mother's ability to conceptualize personality differences, using adjectives along a continuum such as: "Interesting-dull" and "out-going-shy", among types of persons would correlate with her young child's school adjustment. Sample personality "types" included: "Someone you like very much," and "Someone you want to help."

In addition to these instruments, all our interviewers wrote brief summaries on each family. Analysis of this information, however, will not be included in this report.

The preceding instruments, with the exception of the parent interview were administered during the fall of 1965, to a subsample of parents whose children had by then entered Kindergarten.

As indicated, the open-ended Parent Interview schedule was constructed for assessment of parental variables. Considering the number of parent interview schedules now available, it is worthwhile to present some of the rationale behind the construction of another now. Many of the current interview schedules were designed for use with middle or upper-middle class populations. Adaptation of some of the items to lower-class populations, therefore, had been or would have been accomplished only with some time and difficulty, since we cannot assume that the response of a lower class mother to an item reflects the same understanding of its meaning as that of a middle class mother. The revisional process, as a result, would probably have been quite time-consuming with no immediate guarantee of the desired results.

Furthermore, the theoretical assumptions behind some of the instruments now available were not thought necessarily strictly applicable to the present population. The role of the nuclear-family in the lower-class child's growth and development is likely, for example, to be somewhat different from that of the middle class child, since the former comes from a background in which the extended family pattern is a very frequent normative occurance. In the past more emphasis has been given to the influence of familial patterns on psycho-sexual development than to the role of cognitive environmental factors in children's intellectual development. Our goal was to describe the home environments of these children as they were reflected in their mothers, so as to depict graphically their daily experiences from more than an impressionistic or case study viewpoint.

A copy of our interview schedule is not included with this report but is available upon request (D. Slaughter, Urban Child Center, 5801 S. Kenwood, Chicago). Generally, the parental variables studied included measures of the mothers' perceptions of their children's personalities; their achievement expectation--social and academic--for these children; their own self-images as individuals; and their personal expectations as to daily life experiences and future plans. Within this general framework, however, there was considerable variation among interview items. Sample questions included those which were designed to determine just how the mother conceptualized her daily life--how she thought about it; what her expectations for her child's behavior were and what relation this bore to how much power she may perceive she has to implement these expectations and demands; what information from her child she typically registered and responded to, and so forth. Also included were some items which might be specifically predictive of behavior problems with the children such as prevalence of feeding or sleeping disturbances. In addition, each interviewer was instructed to make three specific comments on 1) the events leading up to each interview (an index of any preliminary problems possibly affecting rapport) 2) the family life situation, and personalities of the mother and child as seen by the interviewer, and 3) the appearance of the family's living room as rated on a nine-point scale. Finally, each parent was asked directly about 1) the number of different areas the child is permitted to use for play purposes and 2) the number of available labor-aaving appliances to the mother.

Our approach to this research was through the Head Start Centers in which our Head Start children were serviced. Each parent received a letter indicating that in conjunction with the summer's program, research was being conducted to: 1) assess the needs and behaviors of the children this summer and 2) thereby plan for future Head Start programs. There were informed that to facilitate this planning and assessment, extensive parent interviews would also be conducted, and their cooperation was encouraged.

As a result, during the summer of 1965, 253 mothers of Head Start children in the Chicago area were interviewed by trained interviewers. Each interview took approximately 2½ hours to complete. The results of 228 of these mothers are included in this report.

Most of the interviewers were female teachers especially trained for work with this project. Their teaching experience proved to be quite beneficial: they were accustomed to meeting and discussing children with parents and they were extremely enthusiastic about an opportunity to do so at greater length during the summer. Since most of our Head Start sample was Negro, it was thought that to facilitate rapport the interviewers should be also. Consequently, nine of the eleven interviewers were female, though only two were non-Negro (they were ex-social workers.). The other two interviewers were males, also non-Negro. In this research, however, we found that neither the race nor the sex of the interviewer significantly influenced the receptivity of the respondents, and consequently, the quality of the interviews. Of most importance was the technique of the individual interviewer, particularly the approach he or she used to establish rapport and the personal conviction held as to the importance or significance of the research. Certainly the low refusal rate we received attests to the validity of this approach (see Table VI-1 below).

TABLE VI - 1

TABULATION OF CONTACTS WITH PARENTS

	Center A	Center B
Total number H.S. children enrolled	153	104
Parent Interviews completed	144	84
Refusals	1	1
Number of sets of Twins (mother interviewed for only one child)	2	6
No contact made	5	6
Child Dropped out of program so mother not interviewed	1	3

C. Measures

1. Variables

For purposes of this analysis, we selected twenty-five maternal variables. Eleven of the 25 variables correlated with IQ were derived from the parent interview, 3 from the Education Attitude Survey, 8 from the Behavior Inventory, and the remainder from the Draw-A-Person, Role Repetory, and Embedded Figures Test respectively. A more detailed description of the measures of five of the parent interview variables is included below.

a) Frequency of Mother's Inperative Statements.

As defined in this research, an Imperative statement is essentially a command made by the mother to the child to behave in a certain manner, which is absolute, and which is not accompanied by a rationale or explanation of the necessity of such conformity. Commands could be made by the mother with reference to any aspect of her child's behavior, but in this instance, the specific measure was based on her response to the interview question: "How did you prepare your child for his first day in preschool?"

An example of a response with a high frequency of imperative commands was the following:

> "I told her she supposed to mind the teacher./Do what she say do./She not supposed to fight/ and she supposed to cooperate with the other kids."

Missing in this type of statement is any explanation to the child of why, for example, she is supposed to mind the teacher, or how, for example, she is supposed to "cooperate" with the other children. Were this information also included, this would be classified as primarily an instructive, as opposed to imperative, statement.

b) Directness of Response to Children's Questions.

This variable was an overall rating of the mother's response to the question: What do you do when your child asks you a question you don't want to answer? The issue was: how does the mother handle situations in which her child's curiosity challenges or arouses her own anxiety. Below is the measure of this variable, specifically a five-point scale.

1) Denial. The mother reports that she would specifically state "I don't know", or pretend she hadn't heard the child. The appearance she communicates is one of managing the child by essentially "one-upping" him or besting him in what she sees as direct challenge to her authority.

2) Diversion. The mother reports that she would attempt to divert her child's attention by directing him towards other activities, such as playing with his toys. The implication is that she essentially avoids the issue, hoping this way to avoid handling unpleasant material.

3) Delay. The mother reports that she would postpone answering her child, possibly responding by asking him to wait a few minutes for an answer. Again, this may be a way of diverting the child, but this time the mother does not imply by her actions that she will not answer, she simply uses a bit of "psychology", relying on his own shifting interests.

4) Directness. The mother reports that she feels free to respond directly to any question her child might ask of her. Despite this apparent "freedom", however, the mother does not communicate that she realizes she must modify the manner she responds to correspond to the child's own developmental age and experience.

5) Definition. The mother reports that she responds freely to any question her child might ask of her and that she usually attempts to do so in as simple or uncomplicated a manner as possible. The implication is that she self-consciously modulates the feedback she gives her child, because children understand less than adults.

c) Change Aspects of Child's Personality

An attempt was made to measure the degree of the mother's current satisfaction with her child in the question "If you could change anything about your child, what would you like to do?" Mothers' responses were then rated according to whether they would:

1) Change nothing; he is perfect as he is
2) Change nothing; he is acceptable to her as he is
3) Some aspects of her child's personality or appearance would be changed by the mother.
4) Change practically everything; mother finds child almost completely unacceptable as he is now.

d) Fears of the Child, according to Mother

Mothers were asked a number of specific questions regarding the fears of their children, including Darkness, Thunder and Lightening, Hurting himself by Falling, Sirens or other Loud Noises, Dogs, Doctors, or Other Strangers. For each of these categories, mothers' responses were rated from Very Often Fearful =1 to Never Fearful = 3. A total Fear Score was then computed for all items for each child, independently. Every time a mother selected either 1 or 2, she was given 1 point, and these points were accumulated over the 7 questions. Thus a High Score = a large number of fears for any particular child.

e) Number of Child's Independence Behaviors

A number of age-related child behaviors were described to the mothers, and they in turn were asked to indicate what age they expected their own child to be capable of doing them. It was found, however, that of the 9 items included, many of the mothers stated that their own child was already doing the behaviors. One point was given to each of these items and an Independence Score obtained. The items included:

When do you think your child will be old enough to:

1) Dress or undress himself completely on his own?
2) Pick up his own toys and take care of them?
3) Make his own breakfast himself?
4) Do regular tasks around your house?
5) Make friends with and play with other kids completely on his own?
6) Read stories alone without your help?
7) Take part in your adult interests and conversations with friends?
8) Earn his own spending money?
9) Settle by himself an argument with an older brother or sister?

In addition to the preceding measures mothers' hoped-for and actual expectations of their children were evaluated by asking first "How far do you want your child to go in school?" and then, "How far do you really expect him to go?" A difference score, in terms of actual grade levels mentioned, was computed, although it will not be used in this report, as response to the latter question appears to work equally well as a predictor of children's performances. Finally, frequency of mother's church attendance, ranging from more than once weekly, to less than once a week was also included as a measure of her extrafamilial contacts.

2. Reliability

Preliminary training in the administration of all instruments was given to each interviewer, to insure that uniform procedures would be used. As an added check, each one was asked to describe in some detail events preceding the actual interview, and their own reactions to a particular family so that material obtained under extremely unusual circumstances would not be included in the data analysis. An extensive coding system was devised by this researcher for the parent interview, and coders were trained to utilize it. Depending upon the particular question, intercoder reliability ranged from .87 to 1.00. In the development of the actual interview, every effort was made to reduce the ambiguity and vagueness of questions to mothers. Finally, precoded materials were checked before data processing for any obvious errors in recording.

D. Data Analyses

1. Demographis Characteristics of the Sample

Regarding Table VI-2, generally, the average Head Start child in our sample came from a family whose total family income ranged from $3500-4499 per annum, and whose average monthly rental payment was from $90-104.00. Other data also indicate that 38.8% of the families had received some type of public assistance in the past 5 years. Center B had received the most assistance: approximately 50% of the respondents from Center B had received or were receiving aid. Center A respondents reported that only 30% were on public assistance. The average educational level of the parents was between 9 and 11 years (High School not completed).

2. Results of Two Principal Component Analyses

Since the timeliness of this report was considered an important factor in its presentation, the maternal variables selected for data analysis and included in this report are only a small portion of those available. There, however, seemed the most promising when we asked the question of whether mothers' behaviors would be predictive of Head Start children's academic performance. In all, twenty-five maternal variables from a pool of approximately 150 will be included in this report.

Three of the 25 variables were summary scores obtained from a Principal Component Factor Analysis performed on the Education Attitude Survey, administered to a total of 110 mothers of Head Start children in the Fall of 1965. An additional three variables were obtained from a factor analysis of the twenty socio-emotional items of the Behavior Inventory, also administered to approximately 110 mothers.

TABLE VI - 2

DEMOGRAPHIC CHARACTERISTICS OF THE TWO HEAD START CENTERS

	Center A	Center B
Average Educational Level Completed of Principal wage earner	3.35 (9-11 years)	3.53 (9-11 years)
Average Total Family Income Per Annum	5.30 ($4500-5499)	4.38 ($3500-4499)
Average Monthly Rental or Mortgage Payment of Family	4.70* ($105-119)	3.40 ($90-104)

t $\geq$ .01 Level of Significance

* Center A vs. Center B

3. Results of Two Factor Analyses:

a) Education Attitude Survey

The first factor obtained -- Futility -- appeared to be related to the mothers perception of her own potency regarding the school system. The items loading on the rotated factor 1 included:

Loading on Rotated Factor 1	
.642	If I disagree with the principal there is very little I can do.
.626	I can do very little to improve the schools.
.454	Most children have to be made to learn
.427	Most teachers probably like quiet children better than active ones.

The second factor used - Traditional Educational Values - appeared to be related to the mother's concurrance with normative or typical educational values and attitudes toward school children and the school.

Loading on Rotated Factor 2	
.662	Kids cut up so much that teachers can't teach.
.634	Not enough time is spent learning reading, writing, and arithmetic
.550	When children do not work hard in school, the parents are to blame
.547	Sports and games take up too much time.

The third factor used - Negative Attitudes to Teachers and Curriculum - appeared to denote the degree of the mother's pessisism regarding teachers and the educational system in general.

Loading on Rotated Factor 3	
.748	Most teachers do not want to be bothered by parents coming to see them.
.572	What they teach the kids is out of date.
-.402	Most teachers would be good examples for my children.
-.383	The only way that poor people can raise the way they live is to get a good education.

b) Behavior Inventory

The first factor obtained was a combination of Aggression and Low Achievement items. They are possibly related to the mother's perception of her child's tractibility, possibly both at home and in the classroom.

Aggression and Low Achievement

Loading on Rotated Factor I

.676	Often quarrels with other children his age
.639	When faced with difficult task he either does not try it or gives it up very quickly
.569	Responds to frustration or disappointment by becoming overly angered or bad-tempered
.521	Seems disinterested in how well he does in school or how well he does something I tell him
.510	Insists on maintaining his rights, e.g. insists on getting his turn in group games........
.468	Has little respect for the rights of other chil ren, frequently refuses to wait his turn.

The second factor used, Independent Assertiveness, was apparently related to the mother's perception of her child's ability to assert and use his own inner resources effectively.

Independent Assertiveness

Loading on Rotated Factor 2

.706	Shows a lot of imagination and creativity in the way he plays with his toys.
.690	Appears to trust in his own abilities.
.558	Asks many questions for information about things and people, for example......
.556	Does not need attention or approval from adults to keep on working or playing well
.483	Tries to figure out things for himself before asking adults or other children to help.

The third factor used, Apathy-Sociability, was apparently related to the mother's perception of her child's investment and skills in social situations.

Apathy - Sociability

Loading on Rotated Factor 3	
-.686	Likes to talk or socialize with my adult friends
.594	Often keeps his distance from others because he is shy, suspicious, or disinterested
-.580	Talks eagerly to adults about his own experiences and what he thinks
.560	Has to be pushed before joining others in doing things like playing games, cooperating...
.441	Is slow or apathetic; has little energy or ambition.
-.418	Is eager to tell other children of the things he has done, experiences he has had.

c) Maternal Variables vs. Binet I.Q.

For purposes of this data analysis, it was decided to select those measures of these children's achievement which proved to be the best predictors of their report card evaluations on academic performance; the Preschool Achievement Test correlated .76; and Teacher-Observer ratings at the conclusion of the summer on probable achievement correlated .63 and .56 respectively. Each of these correlations were significant at the .02 level. As a result of these findings, the maternal variables will be examined, using these measures as criteria of children's achievement.

For a total sample of 221 Head Start children of lower class background the mean intelligence level as measured by the Stanford Binet was 90.78, with a standard deviation of 14.51 scale points. Following in Table VI-3 is an indication of the level and direction of representative maternal variables and these Stanford-Binet performance scores.

In interpreting these data, it is to be noted that Head Start children's performances on the Stanford-Binet correlated .62 with the psychologist's rating of their achievement motivation for Binet testing; .62 with Fall Kindergarten teacher's estimates of these children's adaptation; .71 with a National Achievement Test (percentile) administered in Spring, 1966; .66 with a similar estimate of Number readinese; .49 with summer scores on the Draw-A-Person administered on the fourth week of Head Start to each child; and -.41 with summer Head Start teacher predictions of these childrens eventual success (High score - Low success) in Fall Kindergarten. In brief, relative to lower class children's school achievement,

TABLE VI - 3

INTERCORRELATION OF REPRESENTATIVE MATERNAL VARIABLES WITH STANFORD-BINET INTELLIGENCE SCORES OF WORKING CLASS HEAD START CHILDREN

Maternal Variable	Correlation	N	Level of Significance	High
1. Mother's Educational Level (grade completed)	.23	(172)	.01	More educated
2. Number of Independent Behaviors (currently reported by mothers as done by child now	.27	(166)	.01	Very independent
3. Number of Items Feared by child, according to mother	-.16	(152)	.02	Very fearful
4. Level of Educational Aspiration Reported by Mother for Her Child (grade level)	-.23	(93)	.05	Low Aspiration Level
5. Actual Expected Educational Level to be attained by Child, Reported by Mother	-.34	(55)	.02	Low Expected Educational Level
6. Degree of Mother's Desire to Change Aspects of Her Child's Personality	.04	(138)	not significant	Change A Lot
7. Level of Mothers Freedom in Answering Her Child's Questions Directly	.18	(146)	.05	High Freedom
8. Number of Imperatives Used by Mother as Given in First Day Preschool Preparation	-.18	(161)	.01	High Number Imperatives
9. Number of Different Descriptive Categories used By the Mother in Depicting Her Child.	-.05	(142)	not significant	High number categories
10. Frequency of Church Attendence	.07	(150)	not significant	more than once a week
11. Readiness of Child for Kindergarten as reported by Mother on Behavior Inventory	-.26	(85)	.02	Unready

Maternal Variable	Correlation	N	Level of Significance	High
12. Level of Motor Coordination of Child as Rated by Mother on Behavior Inventory	.26	(85)	.02	High coordination
13. Level of Speech Development as rated by Mother in Behavior Inventory	.16	(85)		
14. Level of Overall Adaptation to Kindergarten as Predicted by Mother in Behavior Inventory	-.25	(85)		Low adaptation
15. Level of Academic Achievement in Kinderg. as predicted by Mother on Behavior Inventory	.27	(85)	.02	
16. Degree of Independent Assertiveness as Rated on Behav. Inv.	-.27	(83)	.02	Low Independence
17. Degree of Aggression and Low Achievement rated on Behav. Inv.	.20	(82)	.05	Low Aggression
18. Degree of Apathy-Sociability of Child as Rated by Mother on Behavior Inventory	.24	(82)	.05	
19. Degree of Mother's Expression of Futility on School Situation (Educat. Attitude Survey)	-.12	(84)	not significant	high futility
20. Degree of Mother's Concurrance with Traditional Educational Values (Ed. Att. Survey)	-.04	(83)	Not significant	low agreement
21. Degree of Mother's Negativism Toward Teachers and School Curriculum. (Educ. Att. Survey)	.06	(82)	Not significant	
22. Mothers Intelligence as Measured from Draw-a-Person	-.13	(82)	not significant	High Elaboration
23. Mothers Number of Correct Finds on Embedded Fig. Test	.07	(72)	not significant	High # correct
24. Mothers % of Agreement on Bieri's Role Repetory Test	-.03	(84)	not significant	Low differentiation of roles

Teachers' predictions and test performance scores are considerably better prognostic indicators than mothers' reports. This, of course, was to be expected. Apparently, however, the influence of the home environment upon these children's development is considerably more subtle than these gross measures indicate.

This leads to a second observation: certain kinds of gross measures of maternal variables are more appropriate predictors of children's achievement than others. Mothers do a better job of predicting their child's performance when they are either asked directly about their aspirations and expectations of their child, relative to the school situation or asked about their child's behavior in concrete terms which they observe and relate to on a daily basis. Consider, for example, feeding. A mother of 7th grade educational level responded to the question "How does your child's personality differ from that of her brothers and sisters?" with: "How does she do what?....All of them about the same to me." However, this same mother answered the question "Does your child have trouble eating?" with: "Well, she is kind of a picky eater...she loves hamburger. She don't like chicken. She'll eat an egg if it's boiled. She's kind of a funny eater... The bigger she get, the faster she eat." Clearly, where questions are phrased and directed in the mother's terms it is possible to obtain (from a lower class mother) considerable information regarding her child. In this instance, the mother's attitude toward differences in her children might have been assessed by using as a model of comparison their divergent eating habits.

Similarly, mothers are freer in characterizing children at this preschool age in terms of their speech habits, motor behavior, as contrasted with more direct intellectualized descriptions of their "personalities" and "attitudes." Furthermore, as predictors of their children's performance, measures of mothers' cognitive performance are not terribly useful, generally because of the restricted range and variance of their scores.

Two maternal variables do suggest promise for future researchers: (1) The frequency of Imperatives used by the Mother and (2) the Freedom of Mother to Answer Directly Child's Questions. The frequency of imperatives used by the mother is related to these childrens' fall teachers' ratings of the sophistication of their speech (-.23, N = 78), and overall engagement in Kindergarten (-.20, N = 79); their independence (.27, N = 159, High = low independence) as rated in late summer by Head Start teachers; their achievement motivation as evaluated by the Binet testers (-.24, N = 129, High = low achievement motivation); their mothers affiliation for traditional educational values (.22, N = 80), and their mothers' rating of them as having a higher tendency toward aggressive behaviors (.22, N = 80).

The Directness variable correlated .24 with the number of years of education of the Head Start mother (N = 160); -.24 with her expressions of futility (high) regarding the educational system (N=76); -.24 for both Teacher and Observers' late summer ratings of childrens independence as demonstrated in the nursery (N = 115 and 118 respectively); .28 with these children's performance on the fall Preschool Achievement Test scores (N = 74); and -.22 with children's Impulsivity (N=145) as measured by their ability to postpone acceptance of a smaller piece of candy for a larger one the following day (= delayed gratification). Each of the preceding correlations was significant at the .05 level or better.

d) Maternal Variables vs. PAT Gain Scores

Table VI-4 describes the area in which we obtained some of the higher correlation between maternal and child variables. Mothers' cognitive performance on the Embedded Figures and Role Repetory Tests were predictive of their children's potential for improvement in performance on B. Caldwell's Preschool Achievement Test. More specifically, mothers who were unable to remain patient and find the hidden figures, and who did not make significant distinctions between persons who fulfilled different roles, had children who had potential for making the most improvement over the summer. Interestingly enough, these same mothers were least likely to attend church which, it is well known, is frequently the center of much organized extrafamilial activity within the working class.

TABLE VI - 4

INTERRELATION BETWEEN PRESCHOOL ACHIEVEMENT TEST GAIN SCORE
(PRE AND POST HEAD START) AND SELECTED MATERNAL VARIABLES

Variable	Correlation	N	High
1. Embedded Figures	-.40*	(29)	High=no figures found
2. Role Repetory	.41*	(29)	High Agreement in Description of Different Roles
3. Draw-A-Person	.07	(29)	High elaboration of figures
4. Frequency of Church Attendance	-.33**	(59)	High Attendance

* = significant at .05 or better
** = significant at .02 or better

e) Mother-Teacher-Observer Behavior Inventory Factors

Table VI-5 presents the correlations between mothers' ratings of their children on selected behavioral variables and those of teachers and observers which are comparable (underlined correlations on Table). Clearly, mothers tended to use different criteria in judging or evaluating their children on these dimensions. A task of future researchers might be to clarify these distinctions. They probably reflect, in large part, these childrens' differential adaptation to the school setting. For exmaple, teachers' achievement ratings correlated .52 with children's Preschool Achievement Test scores, while mothers' achievement ratings correlated only .39.

TABLE VI - 5

INTERCORRELATION OF MOTHER-TEACHER-OBSERVER BEHAVIOR INVENTORY FACTORS

	Achievement Teacher 2	Aggeess. Teacher 2	Verb-Soc Teach. 2	Timid Teach 2	Independ. Teacher 2	Achievem. Observer 1	Aggress Obs. 1	Verb-Soc Obs. 1	Timid Obs. 1	Independ Obs. 1
Agg/Low Achievement Mother	.04	.20	.04	-.11	-.12	.07	.21	.04	.01	-.21
Independence Mother	-.20	-.11	.04	.05	.25	-.11	.01	.10	-.10	.04
Sociability-Apathy Mother	.10	-.20	-.34	.24	.08	.21	-.11	-.31	.25	-.04

N = 84 - 87)

E. Conclusion

In summary, data analysis of selected maternal variables with the Stanford-Binet Intelligence level and Preschool Achievement Test score of Head Start children suggest that maternal variables are most useful in predicting potential for changes in their behavior, rather than actual level of performance. The greatest change, however, was generally made in the PAT performance of those children who initially had the lowest scores. Only in this instance are maternal variables better predictors of children's behavior than the prediction of teachers' or previous test performance. For the most part, correlations between maternal variables and children's test performance ranged in the 20's, though with such large sample sizes, frequently these were significant at the .05 or .02 level. It appears that if time is at a premium, maternal variables are not as essential to examine in the prediction of Head Start childrens' achievement level.

From a methodological viewpoint a number of problems were highlighted in the difficulty of obtaining reliable information from lower-class mothers. Mothers give the best information when questions are directed in terms which they concretely observe and use in evaluating their children's behaviors, rather than when they are asked to conceptualize and reformulate their own perspectives to concur with those of the researcher. For example, we obtained more information about whether a mother "personalized" her child by asking her about her child's eating habits, than by asking her directly to "describe" her child, either alone, or by comparison with his other brothers and/or sisters. If parent interviews are not constructed in these terms, other more structured instruments are equally as effective. Other categories which suit the mothers' terms include references to children's conversation, motor behavior, and the like.

Material from one instance in which mothers, teachers, and observers ratings of the same child was presented in this chapter. Generally, the low correlation between mother and teacher/observer ratings on these items from the Behavior Inventory suggests that mothers do not utilize the same criteria in evaluating their children's behavior as are requisite to successful performance in the classroom. Further support for this hypothesis is presented from one particular Behavior Inventory item included by this researcher. Mothers were asked if they had ever "thought of your child in this way before?" In effect, they were asked whether they had ever considered the possibility of rating their child according to the descriptive categories which had been presented to them. This variable had the following correlation with their children's achievement on Report Cards: .22 with performance on school tasks; .40 with a social conformity rating; .27 with a responsibility rating; .41 with a verbal assertion rating; and .34 with an independence rating. In brief, mothers whose children did well in school reported having thought about their child in terms which were congruent with those utilized by the classroom teacher in her evaluations.

Generally, intercorrelation of the maternal variables was low. We found that variables within the parent interview did not correlate more highly than variables measured on the different instruments. In large part, this is a function of our deliberate selection of independent and distinct kinds of tasks so as to ascertain which would give us the most information about these mothers. We found that for a brief, gross estimate of maternal attitudes regarding the school an instrument such as the Education Attitude Survey was most useful, but that from this we could not then generalize as to how successful the child would actually be in school. This appeared to depend upon whether (1) the mother's cognitive orientation to her child's behaviors was similar to that held by teachers, and (2) the extent to which her own performance on a cognitive task such as Witkin's Embedded Figures Test, or James Bieri's Role Repetory test predicted her child's capacity for change with Head Start intervention.

Interestingly, various measures of children's cognitive achievement correlate differently with certain maternal variables. For example, the Frequency of Imperative Commands given by the mothers to the children correlated..47 with their potential for reading readiness, but -.18 with their Stanford-Binet intelligence scores. These two achievement measures, however, correlate .72. Since Binet IQ is the better long-range predictor of school achievement, possibly what these mothers now do is somewhat functional in their development, but has negative repercussions for later learning. Hence, the developmental level of the child might also be relevant to those maternal variables and instruments we select and use to predict school achievement.

In any case, the maternal variables related to the Head Start child's school achievement are extremely subtle and merit more extensive investigation. An interview constructed around interests of the child which the mother can relate to is crucial. Extensive pilot work is essential for the refinement of such a technique. Furthermore, if possible, mother's responses should be corroborated with observations of her behavior within the context of either the home environment or a nursery school setting. We believe that if global rating scales are constructed around these situations those maternal variables crucial for the child's successful development in school might be isolated.

CHAPTER VII

HEAD START AND SUCCESS IN KINDERGARTEN

Abstract: To determine whether the Head Start program has the desired effect upon school adjustment and achievement and which assessments during Head Start are most accurate for predicting success in kindergarten, two types of analysis were carried out. First, 89 children enrolled in Head Start in Location A were compared on a retest during the fall and spring semesters of kindergarten with 82 children from the same social class background, in the same kindergarten classrooms, who had had no Head Start or other nursery school experience. There were no differences between the Head Start and non-Head Start groups on the Preschool Inventory test of achievement, the Percentile Score on a National Test of Reading Readiness, or the Draw-A-Man IQ. There were significant differences, consistently favoring the Head Start group, on teachers' ratings of progress and achievement in kindergarten and on their ratings of the child's motivation to achieve, independence, verbal participation, and lack of timidity. Although these characteristics are not directly related to achievement per se in kindergarten, they indicate the type of accelerated adaptation to the school environment which may be an important precursor of heightened school achievement in the later grades when traditional academic content becomes more important, of heightened motivation to do well in school, and of such variables as lower drop out rates. In the second type of analysis, correlations between six criteria of school success--score on a national reading readiness test and five scores derived from report cards--and measures from the Head Start summer program were used to determine the best predictors of success in kindergarten. For the majority of these criteria, the best predictor was intelligence (achievement or information); the second most adequate predictor was the Head Start teacher or observer's rating of probable readiness, adaptation, and achievement in kindergarten. The simple multiple correlation of Reading Readiness with Binet IQ and Head Start Teacher's Rating of Achievement is .805. In other words, nearly 65% of the variance in reading readiness at the end of kindergarten can be predicted from a knowledge of the child's IQ and his Head Start teacher's prediction of his kindergarten achievement. In spite of some concern among researchers about the use of tests standardized on middle class populations with Head Start children, the Stanford-Binet has proved to be the best predictor of the majority of criteria of success in kindergarten. Other ratings of behavior by observers and teachers also showed high correlations with kindergarten success. Naturalistic observations of behavior and information gained from mothers were relatively poor predictors of kindergarten achievement.

This chapter was prepared by Judith Torney

A. Comparison of Performance in Kindergarten for Children with Head Start Experience and Those without Head Start Experience

The primary aim of the originators and backers of the Head Start Program was that it prepare children, who would otherwise come unprepared, to adapt and achieve better throughout their school careers. Although one cannot hope to see striking results during Kindergarten, if children who have been in the Head Start program are better prepared to enter school, there should be some evidence of this even in the first year of school. Although this project was primarily intended as a methodological one, to develop new tests and assess the usefulness of old tests, some information was gathered which has bearing on the effect of Head Start on children's achievement.

In Location A 89 of the working class children who had been enrolled in Head Start attended three schools and were retested during the fall semester. At the same time they were retested, all of their kindergarten classmates in these same schools were given the same tasks. Of these 169 Non-Head Start children, 82 were chosen to be compared with the Head Start group. These 82 were all working class children who had had no nursery school experience. The mean chronological age of the two groups in months was practically identical (see Table VII-1) (one child who was repeating kindergarten and was therefore a year older than almost all of his classmates was excluded from the Non-Head Start group. With this exception every working class child with no nursery school experience was used.)

Head Start and Non-Head Start groups did not differ in their performance on the National Test of Reading and Number Readiness, administered by the school in the spring semester. (Table VII-1). Neither were there any significant differences in their performance on the Draw a Man IQ test or on the Preschool Achievement Test. Change scores were computed for those children who took the Preschool Achievement both during the summer Head Start period and during the fall retest. The mean gain over this period in the Partial Item Set was 4.38 points (out of a possible total score of 49). As discussed in the chapter on Cognitive Assessment, this gain does not have the same meaning as an increase in IQ would have. An increase in achievement or information, such as this gain of 4.38, is comparable to a gain in mental age on an IQ test. The child has also increased in chronological age during this period, and his IQ could be the same or even slightly lower if the gain in mental age was the same or less than the corresponding gain in chronological age. In other words, this gain of 4.38 points on the PAT probably represents that acquisition of information during this period that would have occurred naturally without the influence of Head Start. This conclusion is supported by the fact that the Head Start and Non-Head Start groups did not differ in their mean scores on the PAT. In summary, from the information available there is no evidence that Head Start prepares children better for reading (in ways assessed by a Reading Readiness test) or teaches them substantial amounts of information, (of the sort measured by the PAT) or increases their intelligence (as measured by the Draw a Man Test.).

Head Start does, however, apparently have a real influence in preparing the child to adapt, adjust, in a general sense better once he reaches kindergarten.

TABLE VII - 1

COMPARISON DURING KINDERGARTEN OF CHILDREN WITH HEAD START EXPERIENCE AND THOSE WITHOUT HEAD START EXPERIENCE - AGE, SCHOOL PERFORMANCE AND COGNITIVE MEASURES

	N	S.E. of Mean	Mean	Diff. H.S. non- H.S.	Significance & Direction
Chronological Age in Months					
Head Start	90	.42	66.72	.51	Not sig.
Non-Head Start	82	.44	66.21		
Reading Readiness National Percentile					
Head Start	73	2.24	22.73	.71	Not sig.
Non-Head Start	70	2.55	23.44		
Draw a Man IQ					
Head Start	89	1.32	77.26	.20	Not sig.
Non-Head Start	82	1.73	77.46		
Pre-School Achievement Test (Highest possible score = 49)					
Head Start	90	.68	31.57	2.00	Not sig.
Non-Head Start	81	.78	29.57		

The differences between mean scores of Head Start and Non-Head Start groups were significant (and all in the direction of better adaptation by Head Start experienced children) on the fall teachers' ratings of the child's Achievement in Kindergarten, Progress since starting Kindergarten, and Progress in Comparison with his Classmates (see Table VII-2).(Fn: Fall teachers who were doing these ratings were not told whether the child had or had not been enrolled in Head Start, although some may have known in the case of certain children.). The fact that it is the ratings of "progress" that show the most substantial differences suggests that children who have had Head Start experience may approach the school situation with similar points of view but by the time of fall retesting (approximately two months after school has statted) they have begun to pull ahead of their classmates in their general achievement and progress in the classroom situation. It may be that if ratings had been conducted in the spring these differences would have been still greater.

In what specific ways are children who have had Head Start experience different from those who have not had this experience? Table VII-3 presents differences in mean scores for Speech and Participation and for Achievement Motivation and Independence. Children coming from Head Start are rated by their kindergarten teachers as significantly less timid, speaking more frequently, and showing more verbal-social participation. There was no difference between Head Start and Non-Head Start on Aggressiveness. Children who had been in the Head Start program are also rated as showing greater achievement motivation (on items from the Behavior Inventory such as "discontinues activity after exerting a minimum of effert" (not rated true) "When faced with a difficult task, either does not attempt it or gives up quickly" (rated not true) and more independent (on items from the Behavior Inventory such as "tries to figure out things for himself before asking adults or children for help" (rated true) and "Does not need attention or approval from adults to sustain him in his work") (rated true). There was no difference between Head Start and Non-Head Start children on the three Summary Score rated by teachers on scales from the Stanford-Binet Face Sheet--Achievement Motivation, Confidence in Ability, Activity Level. It is difficult to account for the difference appearing on one score measuring achievement motivation and not on the other. The items differed somewhat, as discussed more fully in Chapter III.

In summary, Head Start apparently prepares the child to operate within the context of the school situation with greater facility, though it does not produce a specific gain in his IQ, Reading Readiness (as measured on a national test), or information (as measured by the Preschool Achievement Test). Children with Head Start experience are rated as achieving more (in general terms), making more progress,having higher achievement motivation, higher independence, being more independent, more verbal, and less timid than children (from the same type of home backgrounds) who had no nursery school experience. In other words, the child who has been in Head Start is more quickly socialized into the appropriate behavior for the elementary school child. This type of advantage possessed by the Head Start child experience may have real effects in later elementary school years upon drop out rate, disciplinary problems, and school achievement in academic subjects. At the kindergarten level these advantages can only be thought of a precursors of later school achievement and adjustment.

TABLE VII - 2

RATINGS OF BEHAVIOR BY KINDERGARTEN TEACHERS - READINESS, PROGRESS AND GENERAL

	N	S.E. of Mean	Mean	Diff H.S. Non-H.S.	Significance & Direction
Readiness for Kindergarten (Higher scores equal less ready)					
Head Start	90	.11	2.70	.04	Not sig.
Non-Head Start	81	.12	2.74		
Adaptation for Kinderg. (High score = low adapt.)					
Head Start	90	.16	3.70	.41	Not sig.
Non-Head Start	82	.17	4.11		
Achievement in Kinderg. (High score = lower achiev.)					
Head Start	89	.17	3.84	.54	H.S. higher achiev. $p < .02$
Non-Head Start	81	.16	4.38		
Progress Since Starting Kinderg. (high score = less progress)					
Head Start	89	.15	3.01	.62	H.S. more prog. $p < .01$
Non-Head Start	82	.15	3.63		
Comparison of Progress with That of Classmates (High score = greater prog.)					
Head Start	89	.18	4.31	.71	H.S. more prog. $p < .01$
Non-Head Start	82	.18	3.60		
Appearance (High score = cleaner)					
Head Start	89	.12	4.64	.35	Not sig.
Non-Head Start	82	.14	4.99		

(continued on next page)

	N	S.E. of Mean	Mean	Diff H.S. Non-H.S.	Signific. and Direction
Motor Coordination (High score = more coordinated)					
Head Start	90	.13	4.54	.12	Not sig.
Non-Head Start	82	.15	4.43		

TABLE VII - 3

RATINGS OF BEHAVIOR BY KINDERGARTEN TEACHER SPEECH, PARTICIPATION, ACHIEVEMENT MOTIVATION & INDEPENDENCE

	N	S.E. of Mean	Mean	Diff H.S. Non-H.S.	Signific. and Direction
Engagement in School (High score = more engag.)					
Head Start	90	.14	4.58	.42	Not sig.
Non-Head Start	82	.18	4.16		
Speech (High score = more frequent speech)					
Head Start	89	.15	4.51	.75	H.S. more frequ speech $p < .001$
Non-Head Start	82	.19	3.76		
Verbal-Social Participation Sum. Sc. Beh. Inventory (High score = less partic.					
Head Start	86	.17	3.90	.70	Head Start more verb. $p < .01$
Non- Head Start	81	.16	4.60		

	N	S.E.	Mean	Diff H.S. Non-H.S.	Sig. & Direct.
Timidity Sum. Sc.-B.I. (high sc. = less timidity)					
Head Start	86	.14	5.43	.53	H.S. less timid $p < .02$
Non-Head Start	79	.17	4.90		
Aggression Sum Sc. - B.I. (high sc. = less Aggres.)					
Head Start	85	.12	5.85	.17	Not sig.
Non-Head Start	76	.15	6.02		
Achievement Motivation Sum. Sc. - B.I. (high sc. = great. motiv.)					
Head Start	85	.12	5.74	.60	H.S. more Ach. Motiv. $p < .01$
Non-Head Start	81	.18	5.14		
Achievement Motiv. Sum. Sc. - Rated on S.B. Tester Scale (high sc. = gr. ach.)					
Head Start	87	.18	5.49	.09	Not sig
Non-Head Start	81	.21	5.58		
Confidence in Ability Sum. Sc. Rated on S.B. Tester Scale (high sc. = great. gonfid)					
Head Start	83	.18	5.57	.15	Not sig.
Non-Head Start	82	.20	5.42		
Activity Level in Ach. Situation Sum. Sc. Rat. on S.B. Tester Sc.					
Head Start	87	.20	5.35	.00	Hot sig.
Bon-Head Start	82	.22	5.36		

(continued)

	N	S.E.	Mean	Diff H.S. Non-H.S.	Signfi. & Dirct.
Independence Sum. Sc. Beh. I. (high sc. = less Independence)					
Head Start	86	.17	3.57	.87	H.S. more Indep. p < .001
Non-Head Start	81	.18	4.44		

B. Prediction of School Achievement from Information Gathered during Summer Head Start

In addition to determining whether Head Start has any effect upon its participants when they are compared with those having no such experience, it is important to determine whether the children who will have particular problems in Kindergarten can be diagnosed during the Head Start period. If this could be done, perhaps more intensive work could be done with them to alleviate some of these problems. The ability to predict from nursery school experience which children will do well in school also has many advantages in understanding the process of socialization into the school and the types of ability which are necessary for academic success.

1. Defining Variables to be Used

The major criteria chosen for assessing the child's success in Kindergarten were scores on a national test of reading and number readiness--scaled into percentiles and the children's grades on report cards at the end of the fall semester. All these tests and assessments were made as part of the school system's regular program; these were not ratings made for research purposes but the ratings of childrens progress which the teachers sent home to parents and made a permanent part of the school record. The report cards used by this school system are similar to those used to report progress in kindergarten and the early grades in many school systems, including not only progress in Achievement tasks but also various types of social cooperation, discipline, and responsibility that are important in the kindergarten classroom.

Because there were 27 separate ratings, each on a three point scale, on these report cards, the data were factor analyzed to suggest item combinations whic could be used to reduce the number of separate criteria of school success. A Principle Component Analysis with Varimax Rotation of these items was conducted using the population of 84 Head Start children from Location A (see Table VII-4). Six factors were extracted and six summary scores were computed for each child. The first included 4 items, such as "recognizes numerals", "interprets the meaning of Pictures", and is called the Performance of School Tasks. These are ratings which the report card groups under Number. and Reading Readiness. The second factor includes four items and is called the Social Conforming Behavior Score; it includes items such as "respects the rights, opinions, and property of others" and "is kind, polite and thoughtful"; ratings which the report card groups under Social and Emotional Growth. The third score also includes items grouped under Social and Emotional Growth and included five items like "has good self control" and "accepts and carries out responsibility"; this was called the Responsibility Score. The fourth score was called Verbal Assertion and Participation, and included five items, including "contributes to discussion and planning" and "is curious about the world around him." The fifth score included five items, like "experiments with creative material" and "plans and works independently"; this is called the Independence Score. The sixth score included three items concerning Health and muscular control and was not extnesively used. Although the item selection was based upon a factor analysis, these scores were not factor scores. Each Summary Score was the sum of the ratings for the items with the highest loadings on the factor. These items were not weighted according to these factor loadings.

TABLE VII - 4

Summary Score 1: Readiness and Performance of School Tasks[a]

Loading on Rotated Factor 1	Rating and Grouping on Report Card
.676	Recognizes numbers (Number Readiness)
.662	Is learning the meaning of numbers (Number Readiness)
.602	Interprets the meaning of pictures (Reading Readiness
.558	Recognizes likenesses and differences[c] (Reading Readiness)
.446	Is interested in stories, poems, books (Reading Readiness)

Summary Score 2: Socially conforming behavior

Loading on Rotated Factor 2	
.728	Respects right, opinion and property of other (Social and Emotional Growth)
.672	Shares Materials (Soc. and Emot. Growth)
.655	Is kind, polite, thoughtful (Soc. and Emot)
.640	Participates in group activities (Soc & Emot)

Summary Score 3: Responsibility

Loading on Rotated Factor 3	
.662	Has good self control (Soc. & Emot.)
.606	Accepts and carries out responsibilities (Soc & Emot.)
.512	Finishes work on time (Soc. & Emot.)
.500	Works well with others (Work Habits)
.421	Listens to learn; follows directions (Language)

[a] The item "uses own ideas" was loaded .453 on this factor. Its loading on factor 5 was .413 and its content matched the items on this factor more closely than the items on factor 1.

[c] The item "recognizes likeness and diff." was loaded .558 on this factor. However, only 55 children had been given a rating on this and it was therefore not included in the summary score.

Summary Score 4: Verbal Assertion and Participation

Loading on Rotated Factor 4

.718	Contributes to discussion and planning (Language)
.573	Learns and uses new words (Language)
.467	Participates in songs, rhythms, and listening (Arts, crafts, music)
.426	Is curious about the world around him (Science)
.424	Meets new situation and challenges (Soc & Emot)

Summary Score 5: Independence

Loading on Rotated Factor 5

.768	Experiments with creative material (Arts, crafts, music)
.538	Works carefully and neatly (work habits)
.527	Plans and works independently (work Habits)
.456	Works well alone (Work Habits)
.413	Uses own ideas (Soc & Emot.)

Summary Score 6: Health and Physical

Loading on Rotated FActor 6

.734	Is practicing good health habits
.687	Is rested and alert
.610	Is developing muscular control

The factor analysis on which this table is based was a Principal Componant Analysis of the 27 items included in the Report Card in Location A. At the time this analysis was done, only the Head Start sample was available. Varimax Rotation of the 6 factor solution is cited. The number of children graded on these scales ranged from 55 (see note c) to 84. All items were included in a summary score.

The correlation of National Reading Readiness Test with Teacher's Report Card rating of Performance of School Tasks was .803. Because of this high correlation of the two critéria, they are grouped in the following analysis. In considering the Report Card Summary Scores it is important to note that these scores are all correlated with each other significantly. This is one disadvantage of using simple summed scores not factor acores (which by design are independent of each other). The one Report Card Summary Score which was not highly correlated with the others was Social Conformity. There is such a high degree of commonality among our criteria that the discussion will be divided into three parts:Prediction of Reading Standardized Test Score and Prediction of Report Card--Performance of School Tasks; Prediction of Socially Conforming Behavior; Prediction of Responsibility, Verbal Assertion, and Independence.

The variables selected for predicting these criteria of school success were all gathered from the summer testings, and from independent sources such as the mother since the aim was to predict kindergarten success using information gained from sources other than kindergarten teacher. Table VII-5 presents the correlations of selected variables with the Reading Readiness Percentile Score and with the first five Report Card Summary Scores.

2. Predicting Reading Readiness

The best predictor of success in the academic tasks in kindergarten, measured either by score on the reading readiness test or teacher's rating of the Performance of School Tasks, is the Stanford Binet (r = .697 and .762 respectively). Draw a Man IQ was correlated significantly with the two measures of school success, but at a considerably lower level (r = .425 and .390 respectively). The articulation test (# of errors) was nearly as good a predictor as a Draw a Man (r =-.407 and -.358). Neither of the Piaget tasks was a significant predictor. The two measures of low impulsivity showed significant correlations with these variables ranging from .275 to .332.

The second-best predictors of this type of school achievement are specific ratings by either Head Start teachers or observers of Head Start classes of how ready the child is to go to kindergarten, how he will adapt to kindergarten, and how well he will achieve in kindergarten. These correlations were all highly significant and ranged from .459 to .673. There was no tendency for either teachers or observers to be consistently superior in making this type of prediction. These items all came from the Behavior Checklist designed here at the Urban Child Center and differed from items in instruments such as the Behavior Inventory in being directly oriented to prediction of school success. That is to say, teachers and observers were asked to put together all that they know about the child (his intelligence as they have observed it in action, his aggressivity, his speech). This is obviously an important type of question to ask, though it may not be necessary to use all three items since they correlate with each other in the high sixties to low eighties and do not show a great deal of differentiation in their correlation with other items (see Chapter IV).

The third group of variables which predicts Reading Readiness and Report Card Performance of School Tasks are the Summary Scores from the Behavior Inventory as administered to both teachers and observers. When these scores used as predictors, the correlations for teachers are about equal to or slightly better tha those for observers. (Fn: this group of subjects includes only Location A where the original teacher-observer agreement was high; quite different results might have been obtained if information on success had been available for Center B). The correlations with school achievement were highest for the Summary Scores rating Achievement Motivation and lowest using the Summary Score from the Behavior Inventory for Independence.

Because teachers were asked to make these ratings on all children early in the Head Start and again later in the program, it is possible to compare the accuracy of prediction of school success at these two periods. The correlations for a given Summary Score with School Performance for Time 1 and Time 2 are almost identical. In no case (out of 10 comparisons) is a correlation significant at the later time period and insignificant at the earlier time. This suggests that teachers do not need to have extensive experience with children in Head Start in order to make predictions of their success in kindergarten; more precisely, additional weeks of experience do not appear to improve their ability to predict achievement.

In addition to making similar ratings of Probable Achievement and the Behavior Inventory ratings, the most important task the observers performed was making naturalistic observations of time samples of behavior for the Head Start group. These observations were coded, as discussed in Chapter V, and a selection of these variables as predictors of school achievement are presented in Table VII-5 On the whole, these were not very good predictors, probably because of some of the observation difficulties noted in Chapter V. Following each observation period the observers also made precoded ratings of the child's appearance, motor behavior, speech and engagement. Ratings of Engagement, speech and motor coordination were significantly correlated with kindergarten success, but the highest correlation was only .331.

We may consider the most specific task of the observers that of writing down behavior to be coded into categories, next on the continuum of specificity, the precoded ratings following each time sample observation, followed by ratings of specific behavior (as on the Behavior Inventory) where each observer made one rating for the child based upon all the behavior he has observed of that type followed by the least specific rating of how well this child will achieve in kindergarten based on everything the observer has seen the child doing. It appears that the best predictor of school success is that which is least specific within the Head Start situation but most directly oriented toward the kindergarten situation--the prediction of adaptation and achievement.

The children were also rated on their achievement behavior by the Stanford-Binet testers. The drawback to the extensive use of these items is that they are highly correlated with IQ as well as being highly correlated with each other. The correlation of .52 and .619 between Face Sheet Achievement Motivation and school success are almost certainly influenced by the .626 correlation of this variable with IQ. Face Sheet Ratings of Confidence in Ability and Activity are considerably less highly correlated with IQ and show moderate predictive ability.

In the material obtained from mothers, two variables stand out as predictors. The first is the Summary Score from the Mother's adaptation of the Behavior Inventory for Aggression and Low Achievement. The correlation with Reading Readiness does not reach significance, but this variable has a correlation of -.402 with Report Card--Performance of School Tasks. Percentile score on the Readiness Test was correlated .486 with the Number of Imperatives mother gave in the First Day of Preschool task, and her rating of the child's interest or engagement in kindergarten was correlated .315 with Performance of School Tasks. Maternal ratings of child's readiness for kindergarten was not a significant predictor.

In summary, the best predictors of the kind of kindergarten achievement are some measures of the child's intelligence or achievement and the rating by his Head Start teacher or an observer of how well he will achieve or adapt in kindergarten.

3. Prediction of Socially Conforming Behavior

This variable is handled separately from the remainder because it has substantially lower correlations with other report card Summary Scores and lower correlations with predictor variables as well. Its best predictor (r = .40) was the mother's reply to the question: "Have you thought about your child in this way before -- Yes" meaning have you considered his behavior in terms like those used in the Behavior Inventory. It's next best predictors are Binet IQ (.395) and the Head Start teacher's rating of readiness for kindergarten (.388). Other variables show similar patterns of prediction to those reported in the previous section, but all the correlations are appreciably lower. This is apparently a characteristic which is difficult to predict from Head Start.

4. Prediction of Report Card Responsibility, Verbal Assertion, and Independence

For these variables also the best predictors are the cognitive tests of intelligence and achievement. Correlations with the Binet and PAT range from .667 to .720. Impulsivity is a slightly better predictor of this type of behavior adjustment than it was for more academic achievement, and as expected number of errors on the articulation test correlated -.457 with Verbal Assertion.

Moving to the teachers and observers, ratings of readiness, adaptation, and achievement are good predictors (correlations range from .355 to .614) with some sizeable correlations between Behavior Inventory Summary Scores and these non-academic types of kindergarten success. Aggression rated by Head Start teachers shows high negative correlations with the responsibility Summary Score on Report Cards while the Report Card Score on Verbal Assertiveness can be predicted with some accuracy by Head Start Behavior Inventory Ratings of high Verbal-Social Participation, low Timidity, and high Achievement Motivation.

The Naturalistic Observation and the Pre-coded Ratings which followed each time sample show somewhat higher correlations with Verbal Assertion and Independence than they showed with the more achievement oriented variables of Reading Readiness. A number of correlations with the maternal ratings of Readiness for Kindergarten reached significance here as well as the Behavior Inventory Rating made by mothers on Aggression and Low Achievement.

TABLE VII - 5

CORRELATION OF SELECTED VARIABLES FROM SUMMER TESTING WITH SIX CRITERIA OF PERFORMANCE IN KINDERGARTEN - CENTER A, WORKING CLASS ONLY[a]

	%ile Sc.- Nat. Read. Readiness	Summary Scores from Report Cards					Range of N's
		Perform. Sch Task	Social Conform	Responsibility	Verbal Assertion	Independence	
	r	r	r	r	r	r	
Cognitive Variable							
Binet IQ	.717**	.702**	.395**	.562**	.692**	.663**	66-73
Preschool Inventory (PAT)	.697**	.762**	.359**	.567**	.720**	.650**	62-67
Draw A Man (Wk 4)	.425**	.390**	.276**	.347**	.557**	.415**	69-73
Egocentrism (Piag.)	-.253 *	-.196	-.113	-.129	-.063	-.126	66-73
Length Conservation	.141	.015	.100	-.022	.039	-.093	65-73
Impulsivity-Time-Lines	.293**	.327**	.174	.396**	.315**	.403**	65-72
Impulsivity-second of movement	.332**	.275*	.165	.133	.243	.118	55-60
Articulation-Errors	-.407**	-.368**	-.145	-.231	-.457**	-.257*	62-67
By Head Start Teachers							
Readi. for Kinderg.	.528**	.462**	.388 **	.457**	.533**	.414**	67-72
Prob Adapt Kinderg.	.575**	.459**	.337**	.452**	.513**	.454**	67-72
Prob Achiev Kinderg	.627**	.509**	.292**	.422**	.612**	.504**	67-73
Beh I.-Aggress - 1	-.356**	-.178	-.342**	-.410**	-.177	.270**	71-74
Beh I.-Aggress-2	-.353**	-.187	-.282**	-.414**	-.262*	-.282**	69-74
Beh.I.-Verb/Soc-1	.273**	.269**	.266*	.209	.417**	.187	69-74
Beh I.-Verb/Soc-2	.351**	.310**	.271**	.269**	.434**	.189	69-74
Beh I. Timid - 1	-.372**	-.239*	-.272**	-.210	-.467**	-.176	69-74
Beh I. Timid - 2	-.331**	-.211	-.252*	-.160	-.411**	-.091	69-74
Beh I. Indep - 1	.341**	.140	.037	.164	.285**	.312**	69-73
Beh I. Indep - 2	.378**	.226	.115	.259*	.322**	.361**	69-74
Beh I. Achiev - 1	.468**	.366**	.281**	.313**	.493**	.306**	67-72
Beh I. Achiev - 2	.476**	.381**	.231	.349**	.424**	.355**	68-72
By Observers							
Prob Adaptation	.500**	.545**	.316**	.355**	.583**	.448**	52-57
Prob Achievement	.575**	.673**	.312**	.365**	.614**	.560**	52-56
B.I. Aggress-Time 1	-.235*	-.248*	-.218	-.485**	-.202	-.254*	67-72
B.I. Verb/Soc-Time 1	.303**	.371**	.088	.187	.300**	.134	68-73
B.I. Timid Time 1	-.271**	-.252*	-.169	-.181	-.383**	-.139	68-73
B.I. Indep. Time 1	.229*	.120	.188	.266*	.188	.269*	70-73
B.I. Achiev Time 1	.360**	.415**	.328**	.345**	.350**	.375**	68-72

	%ile Sc.-Nat. Read. Readiness	Perform. Sch Task	Social Conform	Responsibility	Verbal Assertion	Independence	Range of N's
	r	r	r	r	r	r	
By Observers (cont.)							
Mean Appear. precod.	.191	.221	.103	.158	.246*	.313*	68-72
Mean Motor precoded	.291**	.237*	-.052	.125	.335**	.333**	68-73
Mead Speech precod.	.215	.315**	-.063	.106	.335**	.182	68-73
Mean Engage. precod	.303**	.331**	.081	.257*	.389**	.353**	68-73
% Aggress Score	-.149	-.179	-.100	-.190	-.030	-.189	69-73
% Auton Ach Score	.030	.031	.206	.116	.148	.044	63-72
% Resistance Score	-.220	-.213	-.299**	-.324**	-.235*	-.357**	68-72
% Unoc & Solitary	-.226	-.253*	.062	-.070	-.084	-.167	68-73
Level Auton Ach	.252*	.169	.216	.190	.233*	.133	68-73
Level Securing Help	.312*	.321*	.057	.351**	.331*	.317*	40-44
By Binet Testers							
Ach Motiv	.520**	.619**	.218	.520**	.601**	.535**	62-68
Confid Abil	.367**	.325**	-.069	.168	.298**	.222	61-67
Activity	.330**	.388**	-.027	.188	.330**	.165	65-71
Information from Mother Interviews							
Readi. for Kindg.	.100	.247	.114	.185	.267*	.278*	54-57
Adapt. for Kindg.	.150	.135	.170	.123	.268*	.248	54-57
Achieve. in Kindg.	.151	.200	.079	.022	.238	.271*	54-57
B.I. Aggr/Low Achiev	-.234	-.402**	-.349**	-.452	-.375**	-.268**	51-54
B.I. Indep	.130	.169	.204	.024	.156	.179	53-56
B.I. Sociab/Apathy	.132	.291*	-.045	.037	.331**	.255	51-55
Engagement	.000	.315**	.083	.151	.207	.201	52-55
Speech	.030	.084	.056	.133	.163	.129	54-57
Motor Coord	.134	.168	-.081	.035	.128	.159	54-57
Appearance	.155	.268*	.085	.143	.304*	.251	54-57
Have thought of child this way	.069	.220	.400**	.272*	.414**	.337**	54-57
# Imperatives	.486**	.037	.023	.106	.255*	.212	65-70
Support child achiev.	.198	.040	.254*	.260*	.082	.088	64-69

a Signs have been changed in the correlations in this Table so that high scores indicate a high amount of the quality named.

CHAPTER VIII

SUMMARY AND CONCLUSIONS

A sudden explosion of interest in measuring the behavior and ability of pre-school children and the assessing of the characteristics of the nursery school situation can be traced to the first Head Start program in the summer of 1965. Although many researchers had studied nursery school children before, the subject of their research had been concentrated in university and upper middle class nursery schools where research was encouraged. Suddenly researchers were face to face with the "culturally deprived" preschooler whom they had reason to believe would become a culturally deprived high school drop-out, unless action programs were instituted to help him in his adjustment to school, beginning before kindergarten.

Although instruments had been developed for preschoolers, many of them required lengthy individual administration and there was considerable question about their interest or adequacy for children from working class backgrounds. This project was designed to make recommendations about the type of instrument which can successfully be used with these children as well as specific suggestions of tests or items from tests which provide useful information about the child, about his Head Start experience, and which may also allow prediction of his success in kindergarten and beyond. It was also intended to make some general statement of the type of methodology which is useful in studying this type of problem.

Because of the assumption that previously used techniques might not be adequate for this population and because of a desire to completely assess all dimensions of a child's behavior, a very large battery of cognitive tests was administered (in one case, as many as five times), extensive naturalistic observations were carried out and carefully coded, teachers were trained to make ratings and fill out extensive checklists and rating scales, and mothers were contacted and given lengthy questionnaires and interviews covering every aspect of child rearing and home environment thought to be important for cognitive and social development. The Head Start groups in Center A and B were an intensively researched collection of pre-school children.

From the conclusions presented in Chapters II to VII those who consider embarking on a similar study can be encouraged. The correlations predicting school success in kindergarten were generally high (multiple correlation of the two best predictors with the criterion of reading readiness was .805). And the data which has been least useful to us may keep other investigators from investigating the same blind alleys.

In general, the most successful techniques for predicting school success were those which bore the closest direct relationship to the type of situation the child will encounter in kindergarten and those which took advantage of the particular skills of the rater or observer without taxing his own cognitive span, For example, the most useful rating made by the teachers for predicting school achievement was their prediction of how well they thought the child would achieve. Teachers know what the school will demand of the child and can assess reasonably well (even close to the beginning of their acquaintance with the Head Start child) how well he will do in kindergarten. Asking more specific questions, such as those in the Behavior Inventory, is important also. If these individual item ratings are combined (both to increase their reliability and increase their range and variance) they give very useful, reliable information about particular characteristics of the child. They present a much more differentiated view of the child, his aggression, his independence, his timidity, than simply a prediction of achievement. This is the sort of variable that would be much more useful in picking out children with special problems and in assessing the impact of particular situations or curriculum. There seem to be real differences in the ability of teachers to predict kindergarten success accurately according to variables such as their experience with preschool and culturally deprived children. This suggests that brief training programs may have to be instituted to train certain teachers to do ratings such as these.

The most highly predictive ratings made by observers were also kindergarten achievement. In every case the more the coders were asked to fit their judgments into a precoded framework in which they felt competent to judge (kindergarten adaptation, motor coordination) the more predictive their judgments. When they were simply asked to write down everything the child did, the task became too diffuse. Though the observations so gathered could be coded fairly reliably, they were not strongly related to other material. The naturalistic observation chapter concludes that observers must be given some kind of framework in which to observe, with some type of precoded behavior ratings and not asked to observe more than one facet of a child's behavior simultaneously (his speech, his expression, his interaction with others). The number of distinctions the observer is asked to make must be reduced to a manageable number. The important relationship between the situation and its demands and the child's behavior brought out in the section on naturalistic observation, suggests that further research should include some categorization of the situation and particularly of the teacher's behavior. Although Centers A and B in this study were very different in environment and personnel, the available information did not allow a complete enough study of the consequences of these differences for the children.

The cognitive domain is, at least superficially, not similar at all to behavior ratings, but here also the type of measure which is closest to the criteria seems to be the best prediction. IQ and achievement tests measure the type of behavior that school success demands--and they are its best predictor. (This is true in spite of the fact that some of these tests were standardized on middle class populations). For particular problem areas, other tests may have value, but are not in most cases substitutes for tests like the Stanford Binet or the Preschool Inventory (PAT).

In interviewing mothers also, it is best to concentrate on those areas of child behavior which are most like those the mother generally thinks about. This is perhaps the area in which the background of working class children enrolled in Head Start presents the greatest need for modification of existing research methods. The characteristics of these mothers which influence their children and which may send them to school "culturally deprived" are precisely those characteristics which are obvious in a gross way but which are difficult to measure, because these mothers can give only limited information about them. In some cases it may be necessary to make comparison of these mothers with middle class mothers to really understand the meaning of maternal variables, OR it may be necessary to discuss problems which are of concern to the mothers, and then rate simple variables like warmth, rather than answers to specific questions. The results of this study do not suggest that the maternal environment of Head Statt children is unimportant, but rather that it is difficult to measure it in ways which can be directly related to school performance or cognitive ability.

With some of the techniques presented here there were measurement problems which have been discussed. Some present instruments may be exceedingly important in a particular school or for a particular investigator which justify further work to refine them. Our major recommendations of instruments and summary scores suggest a baseline of techniques which may be used; others will want to add techniques to measure variables in which they have a particular interest.

Bibliography

1. Ausubel, D. & Ausubel, P., Ego Development Among Segregated Negro Children. In Passow, A. Education in Depressed Areas, New York, Columbia Univ. Press, 1963, pp. 109-141.

2. Bieri, J., Cognitive Complexity: Assessment Issues in the Study of Cognitive Structures. APA paper, delivered in Chicago, Ill., September 4, 1965, Division 16: The Cognitive Approach to Personality Assessment

3. Brim, O., Personality Development as Role-Learning. In Iscoe, I. & Stevenson, H. (Eds.), Personality Development in Children, Austin, Texas, University of Texas Press, 1960

4. Buck, M., Socialization Differences Between Negro ADC Mothers and the Sears, Maccoby, & Levin Norms, Non-published Master's Thesis, Dept. of Sociology, University of Chicago, 1963

5. Caldwell, Bettye M., Preschool Inventory Manual. Unpublished Manuscript, Syracuse, New York, State Univ. of New York, 1965

6. Caldwell, Bettye M. & Soule, D. The Preschool Inventory. Children's Center, Department of Pediatrics, Upstate Medical Center, Syracuse, New York, 1966

7. Crandall, V., Achievement. In National Society for the Study of Education, Pt. I, Stevenson, H. (Ed.), 62: 416-459, 1963

8. Dave, R., The Identification and Measurement of Environmental Process Variables That Are Related to Educational Achievement. Unpub. Ph.D. Disst., Dept. of Education, University of Chicago, 1963

9. Davis, A. Social-Class Influences upon Learning. Cambridge, Harvard Univer. Press, 1948.

10. Dyk, R. & Witkin, H., Family Experiences Related to the Development of Differentiation in Children, Child Development, 36: 21-55, 1965

11. Flavell, J. H. The Developmental Psychology of Jean Piaget. Princeton, New Jersey, D. Van Nostrand, 1963

12. Gildea, M., et. al., Maternal Attitudes and General Adjustment in School Children, In Glidewell, J. (Ed.), Parental Attitudes and Child Behavior, Springfield, Ill., Charles Thomas, Publ, 1961, pp. 42-90.

13. Goodenough, F. L. Measurement of Intelligence by Drawings. New York, World Book, 1926

APPENDIX A-1

BEHAVIOR INVENTORY RATING SHEET[a]

Instructions:

Please rate how this child behaves by circling one of the seven responses to each question.

	Exactly Like	Very Much Like	Quite A Bit Like	Pretty Much Like	Somewhat Like	Very Little Like	Not At All Like
1. Is usually carefree, rarely becomes frightened or apprehensive.	1	2	3	4	5	6	7
2. Talks eagerly to adults about his own experiences and what he thinks.	1	2	3	4	5	6	7
3. Often keeps aloof from others because he is uninterested, suspicious or bashful.	1	2	3	4	5	6	7
4. Tries to figure out things for himself before asking adults or other children for help.	1	2	3	4	5	6	7
5. Has little respect for the rights of other children; refuses to wait his turn, takes toys other children are playing with, etc.	1	2	3	4	5	6	7
6. Seems disinterested in the general quality of his performance.	1	2	3	4	5	6	7

[a] This includes only those items which were administered to kindergarten teachers in the fall retest. These items as well as 14 others were administered in the summer testing to both Headstart teachers and observers. The items adminstered to mothers differed in wording somewhat.

7.	When faced with a difficult task, he either does not attempt it or gives up very quickly.	1	2	3	4	5	6	7
8.	Likes to talk or socialize with teacher.	1	2	3	4	5	6	7
9.	Requires the company of other children; finds it difficult to work or play by self.	1	2	3	4	5	6	7
10.	Is eager to inform other children of the experiences he has had.	1	2	3	4	5	6	7
11.	Appears to trust in his own abilities.	1	2	3	4	5	6	7
12.	Responds to frustration or disappointment by becoming aggressive or enraged.	1	2	3	4	5	6	7
13.	Is constricted, inhibited or timid; needs to be urged before engaging in activities.	1	2	3	4	5	6	7
14.	Asks many questions for information about things, persons, etc. (Emphasis here should be on question prompted by genuine curiosity rather than bids for attention).	1	2	3	4	5	6	7
15.	Emotional response is customarily overstrong; over-responds to usual classroom problems, frustrations and difficulties.	1	2	3	4	5	6	7

[b] Not included in a summary score--see Chapter III. These items could be omitted if the 5 summary scores are to be used.

APPENDIX A-3

16.	Is lethargic or apathetic; has little energy or drive.	1	2	3	4	5	6	7
17.	Is often quarrelsome with classmates for minor reasons.	1	2	3	4	5	6	7
[b]18.	Demonstrates imaginativeness and creativity in his use of toys and play materials.	1	2	3	4	5	6	7
19.	Does not need attention or approval from adults to sustain him in his work or play.	1	2	3	4	5	6	7
20.	Has a tendency to discontinue activities after exerting a minimum of effort.	1	2	3	4	5	6	7
[b]21.	Goes about activities with a minimum of assistance from others.	1	2	3	4	5	6	7
22.	Insists on maintaining his rights, e.g. will not yield his place at painting or at the carpentry bench, etc.; insists on getting his turn on the slide or at group games, etc.	1	2	3	4	5	6	7
23.	Often will not engage in activities unless strongly encouraged.	1	2	3	4	5	6	7
		Much		Some				None
[c]	How much progress has this child made since starting kindergarten this fall.	1	2	3	4	5	6	7
		Below			About the Same			Above
[c]	How does the child's present level of performance compare with his classmates.	1	2	3	4	5	6	7

[c] Included in Fall Retest but not summer testing.

APPENDIX A-4

FACTORS AFFECTING TEST PERFORMANCE

Attention										
1) Absorbed by task	1	2	3	4	5	6	7	8	9[a]	Easily distracted
Reactions During Test Performance										
[b]1) Normal activity level	1	2	3	4	5	6	7	8	9	Hyperactive or depressed
2) Initiates activity	1	2	3	4	5	6	7	8	9	Waits to be told
3) Quick to respond	1	2	3	4	5	6	7	8	9	Urging needed
Emotional Independence										
1) Socially confident	1	2	3	4	5	6	7	8	9	Shy, reserved, reticent
2) Realistically self-confident	1	2	3	4	5	6	7	8	9	Distrusts own ability or overconfident
3) Comfortable in adult company	1	2	3	4	5	6	7	8	9	Ill-at-ease
4) Assured	1	2	3	4	5	6	7	8	9	Anxious about success
Problem Solving Behavior										
1) Persistent	1	2	3	4	5	6	7	8	9	Gives up easily or can't give up
[b]2) Reacts to failure realistically	1	2	3	4	5	6	7	8	9	Withdrawing, hostile, or denying
3) Eager to continue	1	2	3	4	5	6	7	8	9	Seeks to terminate
4) Challenged by hard tasks	1	2	3	4	5	6	7	8	9	Prefers only easy tasks
Independence of Examiner Support										
1) Needs minimum of commendation	1	2	3	4	5	6	7	8	9	Needs constant praise and encouragement

Was is hard to establish a positive relationship with this person?

---

[a] This scaling represents an adaptation of the 10-50 scaling originally presented on the Binet Face Sheet. The 10-50 scaling was used by testers in the summer, the 1 to 9 by kindergarten teachers in the fall.

[b] Not included in a summary score. See Chapter III. These items could be omitted if the 3 summary scores are to be used.

APPENDIX A-5

Child #____________________School____________
Teacher_____________________Date____________

VII. CHECKLIST EVALUATION OF READINESS FOR KINDERGARTEN:[a]

Kindergarten is often the first school experience for many children. Some children may not be ready for kindergarten, while others may be ready in some ways and unprepared in others. Please check the statement below that best describes this child's readiness for Kindergarten.

1. This child was:
   1. ready in all ways for kindergarten.________________
   2. ready in most ways for kindergarten._______________
   3. ready in some though unready in other ways for kindergarten.______________
   4. unready in most ways for kindergarten.______________
   5. entirely unready for kindergarten. Kindergarten should have been postponed for a year. ______________

Social-emotional		(1) Very True	(2) Somewhat true	(3) Not true
2. not able to control his behavior	2.	1	2	3
3. not able to play with other children	3.	1	2	3
4. not able to take guidance from teacher	4.	1	2	3
5. other	5.	1	2	3

[a] The form presented is that used by kindergarten teachers in the fall. Similar questions were used for Headstart teachers in the summer.

APPENDIX A-6

		Very True	Somewhat True	Not True
Physical				
6. not able to control bowels or bladder	6.	1	2	3
7. not able to remain quiet for any period of time.	7.	1	2	3
8. not able to control his bodily movements	8.	1	2	3
Cognitive				
9. not able to speak or communicate his needs well enough to be easily understood.	9.	1	2	3
10. insufficient experience to gain ideas and skills that are needed for kindergarten.	10.	1	2	3

APPENDIX A-7

11. **Appearance:** How does the child look?

Very neat, clean and polished		Very neat and clean		Somewhat unkempt	Very unkempt and dirty	
7	6	5	4	3	2	1

12. **Motor Coordination:** How does the child handle his body?

Excellent balance and coordination		Good balance and coordination		Fairly clumsy and awkward		Very clumsy and awkward
7	6	5	4	3	2	1

13. **Speech:** How often does the child verbalize during school sessions?

Very often		Often		Occassionally		Rarely
7	6	5	4	3	2	1

14. **Engagement:** Please rate on a 7-1 scale your reaction to the child's overall engagement in the kindergarten. You should take into account whether or not the child seems generally able to utilize the opportunities for play, interaction, and learning that surround him. A child rated at 7 would be engrossed, perhaps enthusiastically, in what he was doing while a child at 1 would be withdrawn.

7	6	5	4	3	2	1

APPENDIX A-8

ADAPTATION-ACHIEVEMENT

Keeping in mind the average public school program for the early years, please make the following ratings on this child.

15. Adaptation: We are interested to know how well you think this child will adapt socially and emotionally to the classroom situation in the early school years.

1	2	3	4	5	6	7
Will adapt exceedingly well	Will adapt very well	Will adapt quite well	Will adapt moderately well	Will adapt somewhat well	Will adapt little	Will adapt minimally

16. Achievement: We are interested to know how well you think this child will achieve in the early school years in comparison to his peers.

1	2	3	4	5	6	7
Will excell	Will do very well	Will do quite well	Will be an average student	Will be a little below average	Will not do very well	Will fail

APPENDIX A-9

EGOCENTRISM TEST

Materials: paper house, one side with windows but no door, and one side with door but no windows.

1. Here is a house. Look at it carefully and then I'll ask some questions about it. (Show both sides)

2. (Hold house at eye level so Examiner can only see side with windows and child can only see side with door.)

Does the house have windows?	Yes	No
Does the house have a door?	Yes	No
Do I see the windows now?	Yes	No
Do I see the door now?	Yes	No
Do you see the windows now?	Yes	No

Score correct if:

Says house has windows
Examiner can see windows
Examiner can't see door
Child can't see windows

APPENDIX A-10

DRAW-A-MAN INSTRUCTIONS

Say "I would like you to draw a picture of a person, a whole person. Draw the best picture you can. (If the child doesn't understand the word 'person' say man or woman, boy or girl).

If the child starts to draw only part of a person, do NOT repeat the instructions to draw a whole person.

This test should be done with #2 pencil on 8x12 paper.

DRAW-A-CIRCLE AND DRAW-A-LINE SLOWLY

Procedure: Give the child a pencil and, drawing a vertical line, say to him, "you make one--like this. Make it here." Illustrate once only. Give one trial.

FOLLOWING THE TRIAL, Say, "Now I want you to draw one as slowly as you can; be sure to keep the pencil moving." As soon as he stops moving the pencil, say "Very good, now I want you to draw one more, even slower. Do it as slowly as you can." The slow line trial is ended as the child stops moving the pencil. The examiner should note on the protocol the two lines (of three total lines) which are the slow trials and the time taken for each.

Repeat this with circle. One regular, two slow trials.

IMPULSIVITY PROTOCOL

Have the child turn his chair around facing away from the examiner and say, "Now I would like to see how long you can sit very quietly without moving at all. Just sit and don't move and don't talk. Let's see how long you can sit without moving or talking." (This should be said in a pleasant, quiet tone of voice).

SCORING - use stopwatch

Record the second at which the child makes his first movement, and tally the number of movements that the child makes in the first 30 seconds.

APPENDIX A-11

REVISED PRESCHOOL INVENTORY or PRESCHOOL ACHIEVEMENT TEST (THE PARTIAL ITEM SET)

Begin by asking the child the following questions:

1. What is your name?
2. If child gives first name only, probe for last name. For example, "Johnny what? What's your last name?"
3. When is your birthday? (Score Yes for month or date)
4. Where do you live? (Address, location, etc.)

    I. Point to the following parts of the examiner's body and say, "What's this?"
    II. For all items missed in 5-8, go through again, say, "show me your ______."

5. Finger
6. Eye
7. Elbow
8. Heel

    Now ask the child these questions: "How many ________ do you have?

9. Hands
10. Toes
11. Broken arms

    Now ask, "How many wheels does a __________ have?"

12. Bicycle
13. Rowboat
14. (Hold up piece of paper.) "How many corners does this sheet of paper have?"

    For the next few items take out the box of 12 checkers, all the same color. Give the child the opportunity to manipulate them briefly.

Seeing that all the checkers touch one another and occupy more or less the same area, (all flat on table), put the checkers in two groups in front of the child, as follows and ask (Pointing first to one, then the other):

15. 2 & 8 "Which one has more checkers in it?"

16. 5 & 6 "Which one has more checkers in it?"

17. 6 & 6 "Which one has more checkers in it?"

Take away all but 5 of the checkers. Instruct the child as follows: "Put these checkers next to each other in a line/row." See to it that a half-inch space is made between each two checkers. Give whatever guidance is needed to yield a fairly straight row. Say:

18. "Give me the first one." (Note: Credit first-last in terms of a child's choice; i.e. either end of the row of blocks. All subsequent choices should be consistent with that choice.)

19. "Give me the last one."

Next, line up the checkers in a row, all touching. Take out the two differently colored checkers and stack one on top of the other at one end to make an engine. Say, "Let's pretend this is a train. You know what a train is, don't you? You know it has a lot of cars, one after the other, like this."

20. "What pulls the train?"

21. "What do we call the last car on the freight train?"

Detach the page with the line, triagle, circle, and square drawn on it. Give it to the child. Ask him:

22. "What do we call this?" (Circle)

23. (Line)

24. (Square)

25. (Triangle)

If child cannot name shape, ask him to point to ones missed.

Using the same sheet, say to the child, "Now I'd like you to make some drawings. Make one like this," (and point to):

26. Circle

27. Triangle

Now ask the child to point to "the one which is most like a ___________."

28. Tent or teepee

29. Plate/dish

Take the paper from the child and continue with?

30. "Which is bigger, a ball or a bicycle?"

31. "Which usually goes slower, a car or a bicycle?"

32. "Which is heavier, a brick or a shoe?"

33. "I want you to think of all the things your mother gives you to eat at mealtime, and the things she gives you to eat with. Name all the things you can think of." (If the child says nothing after 10 seconds, say, "you know, like bread and forks." Stop after 30 seconds if the child says nothing. Let him continue if he appears to be still thinking.) Circle the number of items the child names.

Code responses as: 1 (Clear, correct), 2 (Approximation), 3 (Other), 0 (No response)

34. What day is today?

35. What do we call the time of year when it's hottest?

36. What do we call the time of year when it's coldest?

37. If your mother wanted to call up and talk to a friend, what would she use?

38. If you wanted to find a lion, where would you look?

Take out the three cars, red, yellow, and blue; take out the three boxes, black, white and green. Be sure the black box is bottoms up. After each item, make sure all cars and boxes are visible and available; i.e., do not leave a car in a box, etc. Give each instruction only once; make sure the child is looking and listening, and say the words slowly.

39. Put a car on a box.

40. Put a car in a box.

41. Put a car under a box.

42. Put the red car on the black box.

APPENDIX A-14

43. Put the blue car on the green box.

44. Put the yellow car on the little box.

45. Put one car in the middle sized box.

46. Put 3 cars in the big box.

47. Put 2 cars behind the box in the middle.

48. Give everything to me.

49. Ask the child to "untie and tie his shoe." Child can tie shoe.

APPENDIX B-1

COMPARISON OF PERFORMANCE
ON THREE ADMINISTRATIONS OF PRE-SCHOOL INVENTORY BY ITEM**

Item		Summer 1		Summer 2		Retest	
		%	N	%	N	%	N
1. Name		97.8	136	100.0	48	100.0	104
2. Last Name		85.0	127	71.1	45	92.3	104
3. Age		82.2	135	89.4	47		
4. Birthday		47.8	134	57.8	45	32.4	102
5. Address		68.4	133	85.7	49	81.0	100
6. School		63.3	128	79.2	48		
7. Teacher's Name		82.4	125	95.8	48		
8. Childrens' Names in Class-First		3.57 Mean	113	3.57 Mean	46		
9. Childrens' Names in Class-Last		2.12 Mean	34	2.44 Mean	16		
10. What is?	ear	97.8	139	100.0	49		
11.	finger	90.2	139	98.0	49	99.0	105
12.	neck	94.9	139	97.9	49		
13.	back	99.3	139	100.0	49		
14.	eye	100.0	139	98.0	49	100	105
15.	elbow	79.0	138	81.6	49	90.4	104
16.	heel	[illegible]	[illegible]	[illegible]	49	86.4	103
17.	shoulder	84.9	139	91.9	49		
18.	eyebrow	88.5	139	93.8	49		
19.	knee	91.3	138	91.9	49		
20. How many?	eyes	88.5	139	91.8	49		
21.	noses	78.4	139	75.5	49		

APPENDIX B-2

22.	ears	89.2	139	91.8	49		
23.	heads	85.6	139	87.8	49		
24.	feet	84.2	139	93.9	49		
25.	hands	77.0	139	79.6	49	45.7	105
26.	toes	12.3	138	25.0	48	12.4	105
27.	mouths	83.5	139	81.6	49		
28.	necks	84.8	138	83.7	49		
29.	broken arms	72.0	138	67.3	49	41.0	105
30.	# Wheels? car	50.4	139	57.1	49		
31.	bicycle	75.5	139	85.7	49	91.4	105
32.	tricycle	56.1	139	65.3	49		
33.	wheelbarrow	23.7	139	34.7	49		
34.	rowboat	22.8	136	42.9	49	15.2	105
35.	Count to five	4.72 Mean	130	5.09 Mean	46		
36.	What is a corner	51.8	139	77.6	49		
37.	# of corners on paper	44.6	139	63.3	49	50.0	102
38.	Checker groups: 2 & 8-more	90.6	139	95.9	49	92.4	105
39.	5 & 6-more	63.8	138	83.7	49	85.7	105
40.	6 & 6-more	28.1	139	34.7	49	6.7	105
41.	8 & 2-fewer	45.7	138	61.2	49		
42.	Checker row: middle one	63.0	138	69.4	49		
43.	first one	58.0	138	57.1	49	90.5	105
44.	last one	49.3	138	46.9	49	41.9	105
45.	second one	42.4	139	44.9	49		
46.	next-to-last	36.7	139	42.9	49		

APPENDIX B-3

47.	Train: name of first car	32.4	136	55.3	47		
48.	name of last car	23.7	135	34.8	46		
49.	what pulls train (eng. or cab.)	62.1	132	71.4	42	52.6	95
50.	last car? (eng or cab.)	58.8	131	75.6	45	25.0	84
51.	Name: circle	55.1	138	83.6	49	90.5	105
52.	line	44.9	138	75.6	49	86.6	105
53.	square	31.4	137	57.1	49	85.7	105
54.	triangle	75.3	138	67.4	49	91.5	105
55.	Imitate: line	94.6	129	93.7	48		
56.	circle	96.1	129	87.5	48	98.1	105
57.	square	55.5	128	75.0	48		
58.	triangle	44.5	128	40.4	47	50.0	102
59.	Which is like: wheel	87.0	138	83.7	49		
60.	window	87.7	138	91.8	49		
61.	string	77.5	138	77.6	49		
62.	tent or teepee	66.4	137	72.9	48	71.4	105
63.	ice cream cone	42.0	138	33.3	48		
64.	plate/dish	58.0	138	71.4	49	72.4	105
65.	stick	90.5	137	87.8	49		
66.	Bigger: ball or bicycle	79.7	138	77.1	48	78.1	105
67.	tree or flower	84.7	137	87.5	48		
68.	telephone or television	79.7	138	91.7	48		
69.	man or boy	80.4	138	79.2	48		
70.	mosquito or grasshopper	68.4	136	79.2	48		

APPENDIX B-4

71. Bigger: fly or butterfly	86.2	138	87.5	48		
72. Slower: horse or dog	54.0	137	52.1	48		
73. car or bicycle	62.0	137	75.0	48	56.7	104
74. train or rocket	48.9	137	62.5	48		
75. Heavier: butterfly or bird	73.2	138	72.9	48		
76. brick or shoe	78.2	101	87.5	48	87.5	104
77. feather or fork	72.5	138	70.8	48		
78. Close your eyes	94.9	138	100.0	49		
79. Raise your hand	98.6	138	91.8	49		
80. Show teeth	99.3	138	95.9	49		
81. Show fingernails	94.2	138	93.9	49		
82. Wiggle	81.2	138	77.6	49		
83. "Hello" - loud	87.7	138	89.8	49		
84. "Hello" - soft	87.7	138	91.8	49		
85. Stand	97.1	137	98.0	49		
86. Turn around	87.0	138	91.8	49		
87. Face door	91.9	136	95.9	49		
88. Jump	97.8	137	98.0	49		
89. Sit	98.5	136	97.9	48		
90. # of things at mealtime	2.62 Mean	134	2.09 Mean	45	2.47 Mean	105
91. What color: red	81.3	132	79.6	49		
92. yellow	80.3	132	73.4	49		
93. orange	82.6	132	81.6	49		
94. green	76.5	132	83.6	49		
95. blue	73.5	132	71.4	49		

APPENDIX B-5

96.	What color: purple	67.4	132	65.3	49		
97.	brown	78.8	132	75.5	49		
98.	black	84.8	132	79.6	49		
99.	What color is: fire	78.2	133	79.1	48		
100.	grass	80.5	133	81.2	48		
101.	snow	78.2	133	81.3	48		
102.	carrot	61.4	132	52.1	48		
103.	sky	51.6	132	64.6	48		
104.	night	72.0	132	83.3	48		
105.	Direction: saw	65.4	133	72.4	47		
106.	elevator	51.1	131	59.6	47		
107.	ferris wheel	42.4	132	36.9	46		
108.	phonograph record	68.9	132	79.2	48		
109.	waterfall	53.0	132	70.2	47		
110.	When eat breakfast	61.5	130	74.5	47		
111.	What day church	55.8	131	53.2	47		
112.	What's today	19.8	131	47	25.5	89	29.2
113.	What outside when bedtime	80.9	131	85.1	47		
114.	Hottest time of year	29.8	131	45.7	46	26.9	89
115.	Coldest time of year	29.0	131	53.3	45	38.8	98
116.	Time of year now	29.1	131	42.6	47		
117.	What used to call up friend	87.7	130	85.1	47	90.3	103
118.	Where find lion	61.6	130	63.9	47	84.5	103
119.	Where buy gas	90.6	128	85.1	47		

APPENDIX B-6

120.	Where go if sick	93.0	129	91.3	46		
121.	Where find boat	76.0	129	87.2	47		
122.	How find something to read	82.1	128	83.0	47		
123.	Car on box	82.7	133	87.2	47	81.9	105
124.	Car in box	97.0	134	97.8	46	98.1	105
125.	Car under box	86.6	134	87.0	46	97.1	105
126.	Red car on black box	1.50 Mean	133	1.57 Mean	47	1.20 Mean	105
127.	Blue car on green box	1.78 Mean	133	1.64 Mean	47	1.85 Mean	105
128.	Yellow car on little box	1.68 Mean	133	1.55 Mean	47	1.24 Mean	105
129.	One car in middle-sized box	2.08 Mean	131	1.85 Mean	47	2.13 Mean	105
130.	Three cars in big box	2.01 Mean	134	1.28 Mean	47	1.30 Mean	105
131.	Two cars behind middle box	2.16 Mean	134	1.93 Mean	46	1.93 Mean	105
132.	Give all to teacher	1.24 Mean	132	1.20 Mean	45	1.13 Mean	104
133.	Trace pre-drawn line	82.0	128	84.4	45		
134.	Draw line from: bird to wagon	89.8	128	91.1	45		
135.	clock to cake	73.2	127	86.7	45		
136.	dog to boy	71.4	126	84.4	45		
137.	girl to boy	77.6	125	86.7	45		
138.	bird to other bird	77.7	121	82.2	45		
139.	Can put on jacket without help	99.2	118	100.0	44		
140.	Can zip or button jacket	95.7	117	100.0	44		

APPENDIX B-7

141.	Wears shoes	99.1	117	100.0	44		
142.	Can put on shoes	96.6	118	93.2	44		
143.	Can put on correct shoes without help	83.6	116	86.4	44		
144.	Can tie shoes	55.3	114	54.5	44	66.7	81
145.	Can carry out instructions	93.0	114	97.7	44		
146.	Can go home alone	70.9	117	69.0	42		
147.	Knows traffic light	89.3	112	81.8	44		
148.	Can wash hands	99.1	115	100.0	43		
149.	Wash & dry hands & face	99.1	115	100.0	43		
150.	Tells teacher of toilet needs	90.4	115	95.3	43		
151.	Can use bathroom without help	99.1	115	97.7	43		
152.	Doctor does?	88.3	128	91.3	46		
153.	Doctor valence	2.85=Mean	106	2.46=Mean	39		
154.	# words in answer	3.24=Mean	125	3.89=Mean	44		
155.	Policeman does?	100.0	129	100.0	45		
156.	Valence	3.46=Mean	114	3.37=Mean	43		
157.	# words in answer	3.36=Mean	121	3.86=Mean	43		
158.	Dentist does?	79.1	129	91.3	46		
159.	Valence	3.12=Mean	107	3.16=Mean	43		
160.	# words in answer	3.27=Mean	128	3.72=Mean	43		
161.	Teacher does?	81.9	127	91.3	46		
162.	Valence	2.52=Mean	106	2.50=Mean	42		
163.	# words in answer	3.40=Mean	122	3.50=Mean	44		
164.	Father does?	90.6	128	95.6	45		

APPENDIX B-8

165.	Father valence	2.99=Mean	116	3.07=Mean	42
166.	# words in answer	2.72=Mean	125	2.98=Mean	44
167.	Nurse does?	77.2	127	80.0	45
168.	Valence	2.59=Mean	94	2.58=Mean	31
169.	# words in answer	3.19=Mean	126	3.67=Mean	46
170.	Mother does?	92.1	127	95.7	46
171.	Valence	2.66=Mean	116	2.70=Mean	43
172.	# words in answer	2.77=Mean	124	3.09=Mean	45
173.	Soldier does?	80.3	127	91.3	46
174.	Valence	3.41=Mean	102	3.51=Mean	41
175.	# words in answer	2.02=Mean	125	2.04=Mean	45
176.	Number right	7.65=Mean	130	8.26=Mean	46

APPENDIX C-1
BEHAVIOR FORM OBSERVATION SHEET H.S.R.P. JULY 16, 1965 /rbk

Observ.__________ Date__________Time Begun__________ Time Ended__________ Child #______________ Page ______

SETTING	STIMULUS	BEHAVIOR FORM (INCLUDING INTERACTION & OBJECTS	STYLE & EXPRESSION
#			
#			
#			
#			

RATINGS: APPEARANCE: 7-6-5-4-3-2-1 MOTOR COORDINATION: 7-6-5-4-3-2-1 SPEECH: 7-6-5-4-3-2-1 ENGAGEMENT: 7-6-5-4-3-2-1